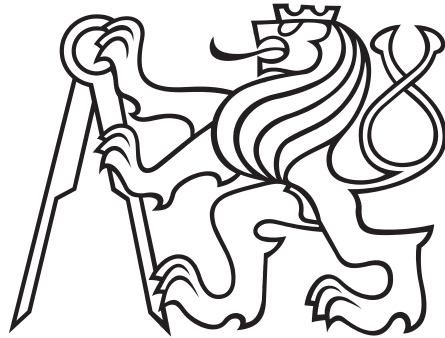


ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
FAKULTA ELEKTROTECHNICKÁ



Predikce krátkovlnného slunečního záření pro malé fotovoltaické systémy

Forecast of shortwave radiation flux for small solar systems

BAKALÁŘSKÁ PRÁCE

Vedoucí práce: Ing. Jan Šulc

Vypracoval: Martin Procházka

Praha 2016

České vysoké učení technické v Praze
Fakulta elektrotechnická

katedra řídicí techniky

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student: **Martin Procházka**

Studijní program: Kybernetika a robotika
Obor: Systémy a řízení

Název tématu: **Predikce krátkovlnného slunečního záření pro malé fotovoltaické systémy**

Pokyny pro vypracování:

1. Proveďte rešerši o dostupných zdrojích pro získání předpovědních dat za účelem predikce dopadajícího krátkovlnného slunečního záření.
2. Statisticky vyhodnoťte úspěšnost a chyby predikce jednotlivých zdrojů s ohledem na reálně změřená data slunečního osvětlení, a to na delším časovém horizontu (alespoň jeden rok).
3. Navrhněte algoritmus pro zpřesnění dostupné predikce, která bude využívat reálně změřených dat slunečního osvětlení.
4. Statisticky vyhodnoťte úspěšnost a zpřesnění navrženého algoritmu. Vyhodnocení proveďte pro několik vybraných míst v České republice.
5. Implementujte tento algoritmus jako službu, která bude předpověď poskytovat v reálném čase.

Seznam odborné literatury:

- [1] Šmídek, J.: Prediktivní řízení malých hybridních solárních systémů. Diplomová práce ČVUT FEL, květen 2015
- [2] Staněk, K.: Fotovoltaika pro budovy, ISBN 978-80-247-4278-6, Grada 2013
- [3] Majer, V. a Hejtmánková, P.: Získání dat pro krátkodobou predikci výroby elektřiny z fotovoltaických elektráren, 11th International Conference ENERGY – ECOLOGY – ECONOMY 2012, Tatranské Matliare, květen 2012

Vedoucí: Ing. Jan Šulc

Platnost zadání: do konce letního semestru 2016/2017

L.S.

prof. Ing. Michael Šebek, DrSc.
vedoucí katedry

prof. Ing. Pavel Ripka, CSc.
děkan

V Praze dne 2. 2. 2016

ABSTRAKT

Hlavním cílem této bakalářské práce je vytvořit samoučící algoritmus pro předpověď intenzity slunečního záření, který využívá měřená data jako zpětnou vazbu pro úpravu původní, volně dostupné predikce. Výstupem tohoto algoritmu je předpověď krátkovlnné radiace pro vybrané lokality v České republice ve W/m^2 na vodorovnou plochu. Součástí práce je rešerše o dostupných zdrojích predikčních a měřených dat, vyhodnocení chyb predikce jednotlivých zdrojů, popis algoritmu a vyhodnocení úspěšnosti výsledné předpovědi.

ABSTRACT

The main goal of this thesis is to create a self-learning algorithm for the solar irradiation prediction, which uses measured data as a feedback to adjust the original, freely available prediction. The output of this algorithm is a prediction of shortwave irradiation for selected locations in the Czech Republic in W/m^2 on horizontal plane. This thesis also consists of research on available sources of prediction and measured data, evaluation of the accuracy of these sources, algorithm description and evaluation of the accuracy of the final prediction.

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, dne

Podpis

Poděkování

Zde bych rád poděkoval panu Ing. Janu Šulcovi za všechny cenné rady, připomínky a pomoc, kterou mi během mé práce na BP poskytl. Dále bych chtěl poděkovat všem členům týmu RP5 Univerzitního centra energeticky efektivních budov ČVUT a všem, kteří mi při práci pomáhali.

OBSAH

1 Úvod.....	9
2 Přehled dostupných zdrojů měřených a předpovědních dat.....	10
2.1 Měřená data.....	10
2.1.1 Univerzitní centrum energeticky efektivních budov.....	10
2.1.2 Stanice Fiedler.....	10
2.1.3 Shrnutí.....	11
2.2 Předpovědní data.....	11
2.2.1 Národní úřad pro oceán a atmosféru.....	11
2.2.1.1 NOAA, GEFS Reforcast v2 Server.....	12
2.2.1.2 ERDDAP NOAA Server.....	12
2.2.2 Solar Radiation Database.....	12
2.2.3 ENERGOCENTRUM PLUS, s.r.o.....	12
2.2.4 Medard.....	13
2.2.5 Pacific Islands Ocean Observing System.....	13
2.2.6 Shrnutí.....	13
2.2.7 Vyhodnocení přesnosti zdrojů.....	14
2.2.7.1 Relative Absolute Error (RAE).....	14
2.2.7.2 RMSE.....	15
2.2.7.3 MBE.....	15
2.2.7.4 MAE.....	16
2.2.7.5 Výsledky.....	16
a) Přesnost predikce na 0-24 hodin.....	16
b) Přesnost predikce na 25-48 hodin.....	17
2.2.7.6 Vytvoření prioritního seznamu.....	18
3 Samoučící algoritmus.....	18
3.1 Obecný popis.....	18
3.2 Funkce pro získání předpovědních a měřených dat.....	19
3.3 Výběr zdroje předpovědních dat.....	20
3.4 Výběr zdroje měřených dat.....	20
3.4.1 Výpočet vzdálenosti dvou lokalit.....	20
3.5 Úprava měřených dat.....	21
3.5.1 Výpočet nekauzálního klouzavého průměru.....	21
3.5.2 Kompensace nízkého počtu zdrojů měřených dat.....	22

3.5.3 Korekce nenulových hodnot v nočních hodinách	23
3.6 Úprava předpovědních dat	23
3.7 Interpolace	24
3.8 Výstupní filtr	24
3.9 Odesílání dat.....	25
3.10 Ochrana proti výpadkům.....	25
4 Výsledky samoučícího algoritmu.....	25
4.1 Dosažená přesnost.....	25
4.1.1 Přesnost predikce na 0-24 hodin	26
4.1.2 Přesnost predikce na 25-48 hodin	27
4.1.3 Přesnost výsledné predikce po měsících	28
4.1.3.1 Predikce na 0-24 hodin.....	28
4.1.3.2 Predikce na 25-48 hodin.....	29
4.2 Příklady predikce.....	30
5 Dostupnost předpovědi.....	30
6 Možnosti ke zlepšení a budoucí práce.....	32
7 Závěr.....	33

1 ÚVOD

Vzhledem ke snižující se ceně bateriových úložišť, zjednodušení administrativního procesu a státním podporám v rámci programu „Zelená úsporám“, počet fotovoltaických instalací s bateriovými úložišti (tzv. hybridních fotovoltaických elektráren) v České republice roste.

Hlavním problémem fotovoltaických elektráren je jejich nerovnoměrná výroba během dne. Maximum výroby se během roku pohybuje kolem poledních hodin. Oproti tomu spotřeba běžného rodinného domu dosahuje maxima v ranních a večerních hodinách. Tento problém dokáže spolehlivě vyřešit bateriové úložiště s patřičným regulátorem. Dle vyhlášky [36] platné pro všechny nové instalace, se musí všechna vyrobená energie spotřebovat lokálně a přetoky do distribuční soustavy jsou zakázány a penalizovány. Jedním z úkolů regulátoru tedy bude zabránit těmto přetokům.

Řízení tohoto hybridního systému, využívající predikci osvitů, pomůže zvýšit finanční úsporu celého systému. Výsledný řídicí algoritmus, v závislosti na predikci, řídí nabíjení akumulátoru nejen z fotovoltaických panelů, ale i z distribuční soustavy (využití nízkého tarifu v nočních hodinách).

Stěžejním výsledkem této bakalářské práce je samoučící algoritmus, který zlepšuje dostupnou predikci na základě zpětné vazby v podobě historicky měřených dat. Výstupem tohoto algoritmu je předpověď intenzity krátkovlnného slunečního záření ve W/m^2 na vodorovnou plochu pro vybrané body v České republice. Dalším úkolem této práce je implementovat tento algoritmus jako službu, poskytující předpověď v reálném čase pro lokální instalace hybridních fotovoltaických systémů.

Výsledek této práce, tedy predikce intenzity krátkovlnného slunečního záření, je významný nejen pro malé fotovoltaické elektrárny, ale i pro velkovýrobce a distributory elektrické energie. V případě dostatečně přesné predikce, předpovídající slunný den, by bylo možné snížit výrobu klasických zdrojů elektrické energie (jaderné a tepelné elektrárny) a ušetřit velké množství finančních prostředků.

Práce je rozdělena do sedmi kapitol. Druhá kapitola obsahuje rešerši o dostupných zdrojích predikčních a měřených dat a vyhodnocení jejich přesnosti. Třetí kapitola popisuje samotný samoučící algoritmus. Následující dvě kapitoly se zabývají vyhodnocením přesnosti predikce a dostupností předpovědi v rámci služby PV Forecast [33]. Šestá kapitola obsahuje možnosti ke zlepšení a poslední kapitolou je závěr.

2 PŘEHLED DOSTUPNÝCH ZDROJŮ MĚŘENÝCH A PŘEDPOVĚDNÍCH DAT

Navržený algoritmus využívá historických změřených dat slunečního osvětlení na několika místech v České republice a několika zdrojů předpovědních dat.

Cílem je získat co nejspolehlivější zdroj předpovědních dat, a to nejenom z hlediska kvality předpovědních dat, ale také s ohledem na dostupnost predikce (frekvence aktualizace dat, četnost výpadku, počet bodů, pro které je predikce poskytována atd.). Tato kapitola obsahuje přehled dostupných zdrojů reálně změřených dat a předpovědi slunečního záření.

2.1 Měřená data

Měření intenzity slunečního záření je důležitou součástí celého projektu. Měřená data představují zpětnou vazbu, která je využita pro zpřesnění předpovědi. V současné chvíli jsou data získávána ze tří meteorologických stanic – UCEEB, Domanín a Nový Lískovec. Tato data jsou nejprve upravena a poté použita jako jeden ze vstupů samoučícího algoritmu. Všechny zdroje poskytují naměřené hodnoty krátkovlnného záření na vodorovnou plochu v jednotkách W/m^2 .

2.1.1 Univerzitní centrum energeticky efektivních budov

První meteorologická stanice je provozována Univerzitním centrem energeticky efektivních budov ČVUT v Praze (UCEEB). Univerzitní centrum leží na okraji města Buštěhrad cca 5 km východně od Kladna. Jsou zde provozovány celkem 3 meteorologické stanice. Jedna je umístěná na střeše objektu a další dvě jsou umístěny v severovýchodní a jihozápadní části areálu. Tyto meteostanice kromě krátkovlnné radiace měří také: teplotu vzduchu (ve výšce dvou metrů a u povrchu země), teplotu půdy (v hloubce 5 cm), tlak vzduchu, rychlost a směr větru, srážky, dlouhovlnnou radiaci a relativní vzdušnou vlhkost. Sběr dat probíhá od 4.11.2014, data jsou ukládána ve formátu .csv a po zadání přístupových údajů dostupná pomocí webového prohlížeče. [28][4]

Data jsou měřena v minutových intervalech, za jeden den je tedy nasbíráno 1440 hodnot. Pro následné použití v algoritmu je nutné z těchto dat vyfiltrovat pouze 24 hodnot, které odpovídají hodinovým průměrům intenzity slunečního záření. Během dne se v datech vyskytují špičky, které výrazně převyšují průměr blízkých hodnot (viz Graf 1 – UCEEB). Pro filtraci je použit tzv. nekauzální klouzavý průměr, který vznikl jednoduchou úpravou kauzálního klouzavého průměru [5]. Samotný výpočet nekauzálního klouzavého průměru je popsán v kapitole 3.5. Denní průběh v jednotlivých krocích obsahuje Graf 1.

2.1.2 Stanice Fiedler

Firma FIEDLER AMS s.r.o. provozuje meteorologické stanice na několika místech v České republice. Pouze dvě z nich však měří krátkovlnnou radiaci. První stanice se nachází v obci Domanín zhruba 4 km jihozápadně od Třeboně a druhá stanice se nachází v Novém Lískovci, městské části na jihozápadě Brna. Obě stanice kromě intenzity slunečního záření měří také tyto hodnoty: rychlost a směr větru, vlhkost (ve výšce 2 a 0,3 metru), srážky, vlhkost půdy, teplotu (ve výšce 2 a 0,3 metru) a teplotu půdy. Data z obou těchto stanic jsou k dispozici od 7.12.2015 a jsou přístupná pomocí webového prohlížeče měřených dat. Data jsou ukládána ve formátu .csv. [6][7]

Intenzita slunečního záření je měřena v desetiminutových intervalech, tedy celkem 144 hodnot za den. Data tedy musí být upravena podobným způsobem jako data ze zdroje UCEEB. Pro získání hodnoty v celou hodinu je použit průměr z 6 hodnot v okolí dané hodiny.

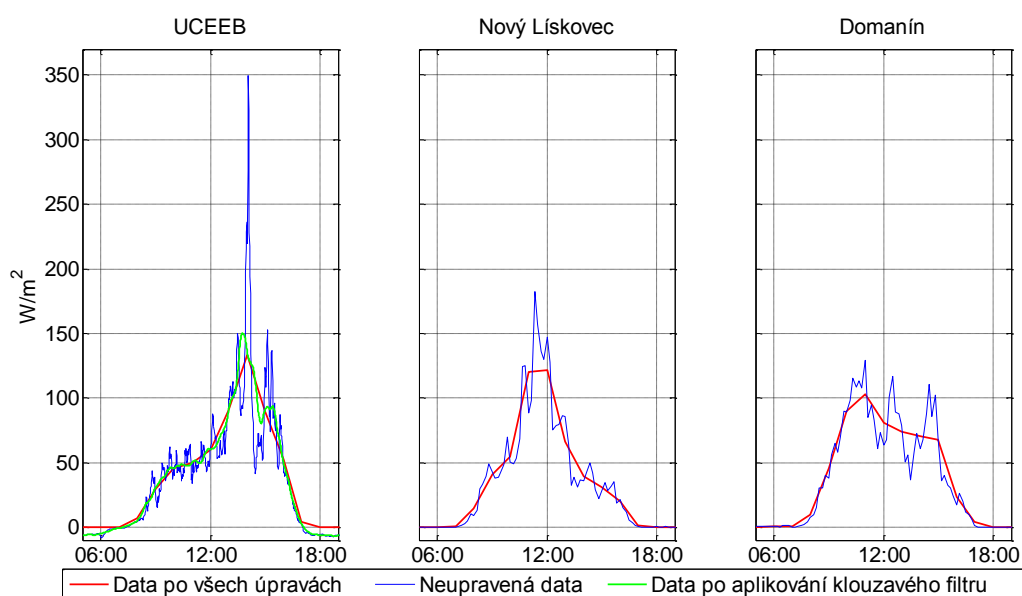
2.1.3 Shrnutí

Informace o všech zdrojích měřených dat se nachází v následující tabulce (Tabulka 1).

Zdroje	Frekvence aktualizace měřených dat [min]	K dispozici od	Pozice GPS
UCEEB	1	4.11.2014	50,2° N, 14,2° E
Domanín	10	7.12.2015	50° N, 14,7° E
Nový Lískovec	10	7.12.2015	49,2° N, 16,5° E

Tabulka 1: Informace o zdrojích měřených dat intenzity slunečního záření.

Příklad hodnot naměřených jednotlivými stanicemi obsahuje Graf 1.



Graf 1: Data naměřená jednotlivými stanicemi dne 1.2.2016.

2.2 Předpovědní data

Navržený algoritmus předpovědi slunečního záření využívá celkem šesti dostupných zdrojů, které poskytují předpověď intenzity slunečního záření v jednotkách W/m^2 na různý horizont predikce a s různou vzorkovací periodou. Tato kapitola se zabývá popisem těchto zdrojů.

Hlavním nedostatkem většiny zdrojů předpovědních dat je „posunutí“ a „roztažení“ předpovědi tak, že nulové hodnoty intenzity záření se v predikci objevují až v cca 23 hodin (i v zimních měsících). Data ze zdrojů trpících tímto problémem jsou upravována ještě před vstupem do algoritmu. Konkrétní provedení úpravy je popsáno v kapitole 3.6.

2.2.1 Národní úřad pro oceán a atmosféru

Národní úřad pro oceán a atmosféru (National Oceanic and Atmospheric Administration – NOAA) je vědecká vládní agentura Ministerstva obchodu Spojených států amerických zaměřená na výzkum v oblasti meteorologie, atmosféry a oceánů. Tato organizace využívá několik různých modelů pro předpověď stavu atmosféry na celé zeměkouli. Data ze dvou těchto

modelů jsou stahována ve formě .grib2 souborů [8]. Data jsou dále zpracována a používána jako zdroj předpovědi krátkovlnného slunečního záření. [9]

2.2.1.1 NOAA, GEFS Reforecast v2 Server

Tento model předpovídá intenzitu slunečního záření na 72 hodin se tříhodinovými intervaly, tedy 8 hodnot na jeden den. Předpověď je dodávána ve formě mřížky s rozlišením jeden úhlový stupeň v zeměpisné šířce i délce. Model předpovídá dalších 111 veličin [12]. Data z tohoto zdroje jsou využívána od 1.12.2015. Pro zjednodušení je tento zdroj dále v textu označován pouze jako „NOAA“. Data z tohoto zdroje mohou být volně použita pod podmínkou uvedení zdroje. [10][11][12]

2.2.1.2 ERDDAP NOAA Server

Tento zdroj dodává předpověď na 48 hodin se tříhodinovými intervaly. Předpověď je aktualizována každých 6 hodin. Předpověď je dodávána ve formě mřížky s rozlišením 0,25, 0,5, nebo 1 úhlový stupeň v zeměpisné šířce i délce. Model předpovídá dalších 164 veličin [13]. Data z tohoto zdroje jsou využívána od 31.12.2015. Pro zjednodušení je tento zdroj dále v textu označován pouze jako „NOAA2“. Data z tohoto zdroje mohou být volně použita a redistribuována. [10][11][13]

2.2.2 Solar Radiation Database

Solar Radiation Database (dále jen „Soda“) je francouzský projekt zabývající se predikcí a dlouhodobým měřením klimatických veličin se zaměřením na veličiny spojené se slunečním zářením. Zdroj Soda poskytuje informace o teplotě vzduchu, relativní vlhkosti, intenzitě slunečního záření a oblačnosti. Data z tohoto zdroje mohou být použita pouze pro osobní účely. [14][15]

Soda dodává předpověď krátkovlnné radiace na 72 hodin se tříhodinovými intervaly, tedy 8 hodnot na jeden den. Předpověď je dodávána ve formě mřížky s rozlišením jeden úhlový stupeň v zeměpisné šířce i délce. Data z tohoto zdroje jsou využívána od 16.12.2015.

2.2.3 ENERGOCENTRUM PLUS, s.r.o.

Firma ENERGOCENTRUM PLUS, s.r.o. (dále jen „Energocentrum“) se zabývá převážně provozováním, servisem a údržbou technických zařízení, měřeními a regulací (MaR), vývojem software a výzkumem v oblasti automatizačních technologií. Společnost dále poskytuje komerční službu pro předpověď počasí. Kromě intenzity slunečního záření (krátkovlnná radiace) dále poskytuje předpověď: teploty vzduchu, oblačnosti, rychlosti větru, srážek, relativní vlhkosti, dlouhovlnné radiace a tlaku. [16][17]

Pro testovací potřeby algoritmu je společností ENERGOCENTRUM PLUS poskytována předpověď intenzity slunečního záření pro jeden bod (Praha). Energocentrum dodává predikci na 72 hodin dopředu s hodinovými intervaly mezi daty, tedy 24 hodnot na jeden den. Data z tohoto zdroje jsou využívána od 18.12.2014.

2.2.4 Medard

Projekt Medard počítá předpověď ve dvou doménách. První doména zahrnuje celou Evropu kromě většiny severovýchodních zemí – Švédsko, Norsko, Finsko a Island. Tato doména má horizontální rozlišení 9 km a předpověď je zde počítána na 3 dny dopředu. Druhá doména s rozlišením 3 km pokrývá území České republiky a předpověď pro ni je počítána na 2 dny dopředu. Pro třetí den je použita předpověď s rozlišením 9 km. Předpověď je projektem Medard aktualizována čtyřikrát denně, vždy v 0, 6, 12 a 18 hodin. Intenzita slunečního záření je predikována na 79 hodin s hodinovými intervaly. [18]

V současné chvíli je pro testovací potřeby algoritmu společností Medard poskytována předpověď pro 3 polohy v České republice, konkrétně polohy, odkud jsou získávána měřená data (viz kapitola 2.1). Data z tohoto zdroje jsou využívána od 9.1.2016.

2.2.5 Pacific Islands Ocean Observing System

Pacific Islands Ocean Observing System (PacIOOS) vznikl vývojem z Hawai‘i Ocean Observing System (HiOOS). HiOOS vznikl spoluprací NOAA a School of Ocean and Earth Science and Technology – University of Hawai‘i at Manoa (SOEST). Tento zdroj je nadále označován pouze jako „Hawaii“. Data z tohoto zdroje mohou být volně použita a redistribuována. [19]

Hawaiská předpověď je dodávána na 120 hodin dopředu s tříhodinovými intervaly a stahována jednou denně. Předpověď je aktualizována ve formě mřížky s rozlišením 0,5 úhlového stupně v zeměpisné šířce i délce. Data z tohoto zdroje jsou využívána od 16.12.2015. Kromě krátkovlnné radiace předpovídá zdroj Hawaii také následující veličiny: teplotu vzduchu (při povrchu a ve 2 metrech), rychlost severního a východního větru v 10 metrech, srážky, relativní vlhkost, tlak na úrovni hladiny moře a dlouhovlnnou radiaci. Data jsou dostupná pomocí webového formuláře (Data Access Form) [20].

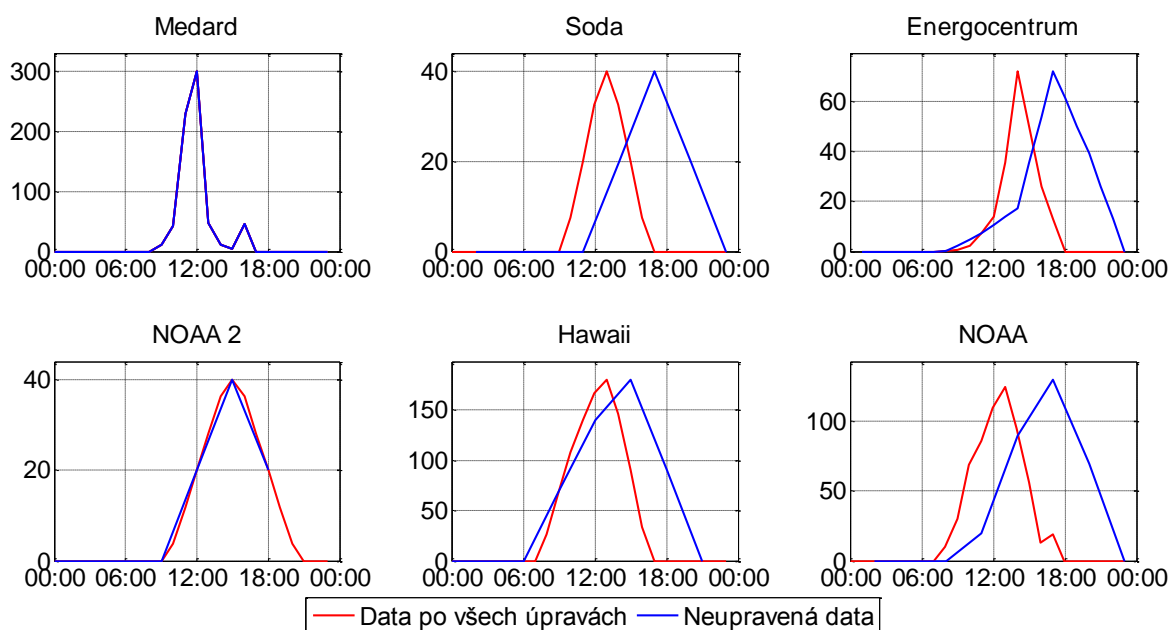
2.2.6 Shrnutí

Informace o všech zdrojích predikce se nachází v následující tabulce (Tabulka 2).

Zdroje	Horizont predikce [h]	Interval mezi daty [h]	Rozlišení	K dispozici od
NOAA	72	3	1°	1.12.2015
NOAA 2	48	3	0,25°,0,5°a 1°	31.12.2015
Soda	72	3	1°	16.12.2015
Energocentrum	72	1	1 bod	18.12.2014
Medard	79	1	3 body	9.1.2016
Hawaii	120	3	0,5°	16.12.2015

Tabulka 2: Informace o zdrojích předpovědních dat.

Příklad hodnot předpovídaných jednotlivými zdroji obsahuje Graf 2. Na grafu je vidět typické posunutí a „roztažení“ předpovědi.



Graf 2: Data předpovídaná jednotlivými zdroji pro den 1.2.2016 a polohu 50° N, 14° E (cca UCEEB).

2.2.7 Vyhodnocení přesnosti zdrojů

Přesnost předpovědních zdrojů byla vyhodnocena několika způsoby. Vyhodnocení bylo provedeno zvlášť pro předpověď na 0-24 a 25-48 hodin dopředu. Způsoby vyhodnocení přesnosti jsou popsány v následujících podkapitolách.

Vyhodnocení přesnosti jednotlivých zdrojů proběhlo v období od 9.1.2016 do 27.4.2016 (109 dní). V tomto intervalu byly již k dispozici všechny zdroje předpovědních a měřených dat.

2.2.7.1 Relative Absolute Error (RAE)

Relative Absolute Error (RAE) neboli relativní absolutní chyba je jedním ze základních způsobů pro výpočet přesnosti předpovídané veličiny.

Přesnost předpovědi v rámci jednoho konkrétního dne byla vyhodnocena pomocí tří různých období vzorce pro výpočet RAE [29][30]:

$$RAE_m = \frac{|s_p - s_m|}{s_m} \cdot 100,$$

$$RAE_t = \frac{|s_p - s_m|}{s_t} \cdot 100,$$

$$RAE_{800} = \frac{|s_p - s_m|}{s_{800}} \cdot 100,$$

kde:

- s_p – denní suma predikovaných hodnot v daný den [Wh/m²],
- s_m – denní suma naměřených hodnot v daný den [Wh/m²],
- s_t – denní suma teoretických hodnot¹ v daný den [Wh/m²],

¹Viz [1] - kapitola 4.1.3 Teoretická hodnota slunečního ozáření

- s_{800} – denní suma teoretických hodnot dne s maximální hodnotou 800 W/m^2 [Wh/m^2],
- RAE_m – relativní absolutní chyba vztažená k denní sumě naměřených hodnot [%],
- RAE_t – relativní absolutní chyba vztažená k denní sumě teoretických hodnot [%],
- RAE_{800} – relativní absolutní chyba vztažená k denní sumě teoretických hodnot dne s maximální hodnotou 800 W/m^2 [%].

Hodnota 800 W/m^2 byla vybrána z toho důvodu, že se jedná o referenční hodnotu, při které jsou testovány solární panely, viz norma ČSN [31].

Tímto způsobem bylo pro každý zdroj získáno 109 hodnot, které odpovídají 109 dnům, na kterých byla přesnost testována. Z těchto hodnot byly poté vypočítány průměry pro jednotlivé zdroje.

2.2.7.2 RMSE

Root Mean Square Error (RMSE) čili střední relativní chyba vztažená k průměru denních sum [29][30]. Výpočet RMSE:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (s_{p,i} - s_{m,i})^2}$$

kde:

- $s_{p,i}$ – denní suma predikovaných hodnot v daný den [Wh/m^2],
- $s_{m,i}$ – denní suma naměřených hodnot v daný den [Wh/m^2],
- N – počet dní, během kterých byla přesnost zdrojů vyhodnocována, $N = 109$ pro náš případ
- $RMSE$ – Root Mean Square Error [Wh/m^2]

2.2.7.3 MBE

Mean Bias Error (MBE) může nabývat záporných i kladných hodnot [29][30]. Výpočet MBE je následující:

$$MBE = \frac{1}{N} \sum_{i=1}^N (s_{p,i} - s_{m,i})$$

kde:

- $s_{p,i}$ – denní suma predikovaných hodnot v daný den [Wh/m^2],
- $s_{m,i}$ – denní suma naměřených hodnot v daný den [Wh/m^2],
- N – počet dní, během kterých byla přesnost zdrojů vyhodnocována, $N = 109$ pro náš případ
- MBE – Mean Bias Error [Wh/m^2]

2.2.7.4 MAE

Mean Absolute Error (MAE) neboli střední absolutní chyba. Tato hodnota může nabývat pouze kladných hodnot [29][30]. Výpočet MAE je následující:

$$MAE = \frac{1}{N} \sum_{i=1}^N |s_{p,i} - s_{m,i}|$$

kde:

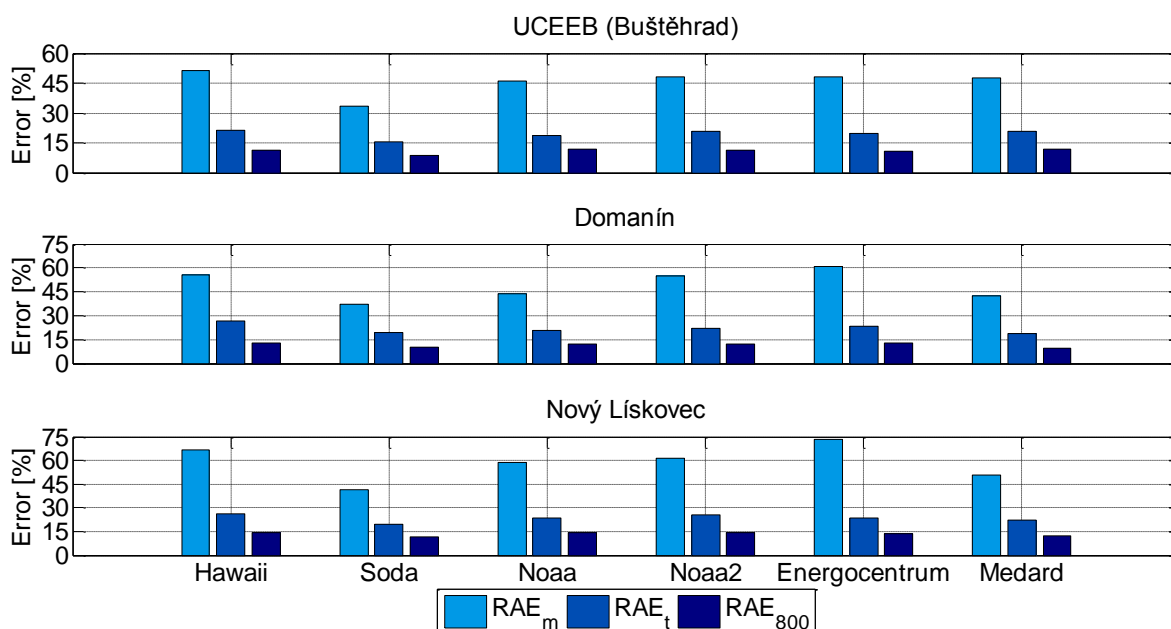
- $s_{p,i}$ – denní suma predikovaných hodnot v daný den [Wh/m^2],
- $s_{m,i}$ – denní suma naměřených hodnot v daný den [Wh/m^2],
- N – počet dní, během kterých byla přesnost zdrojů vyhodnocována, $N = 109$ pro náš případ
- MAE – Mean Absolute Error [Wh/m^2]

2.2.7.5 Výsledky

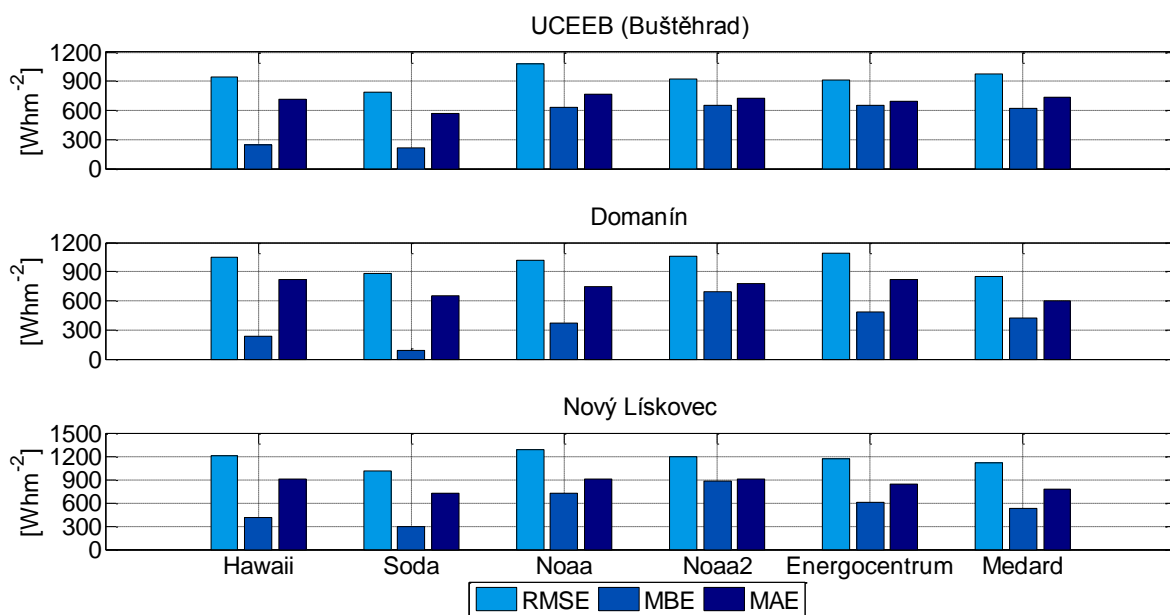
Přesnost predikce byla vyhodnocena zvlášť, a to na 0-24 a 25-48 hodin dopředu. Vyhodnocení přesnosti proběhlo až po úpravě dat způsoby popsány v kapitole 3.6. Přesnost původních předpovědních neupravených dat by byla vzhledem k posunu a „roztážení“ predikce mnohem nižší.

Z následujících grafů je patrné, že nejpřesnějším zdrojem je zdroj Soda, následovaný zdrojem Hawaii. Další v pořadí jsou zdroje Medard a Energocentrum. Nejméně přesné zdroje jsou zdroje NOAA a NOAA2.

a) Přesnost predikce na 0-24 hodin

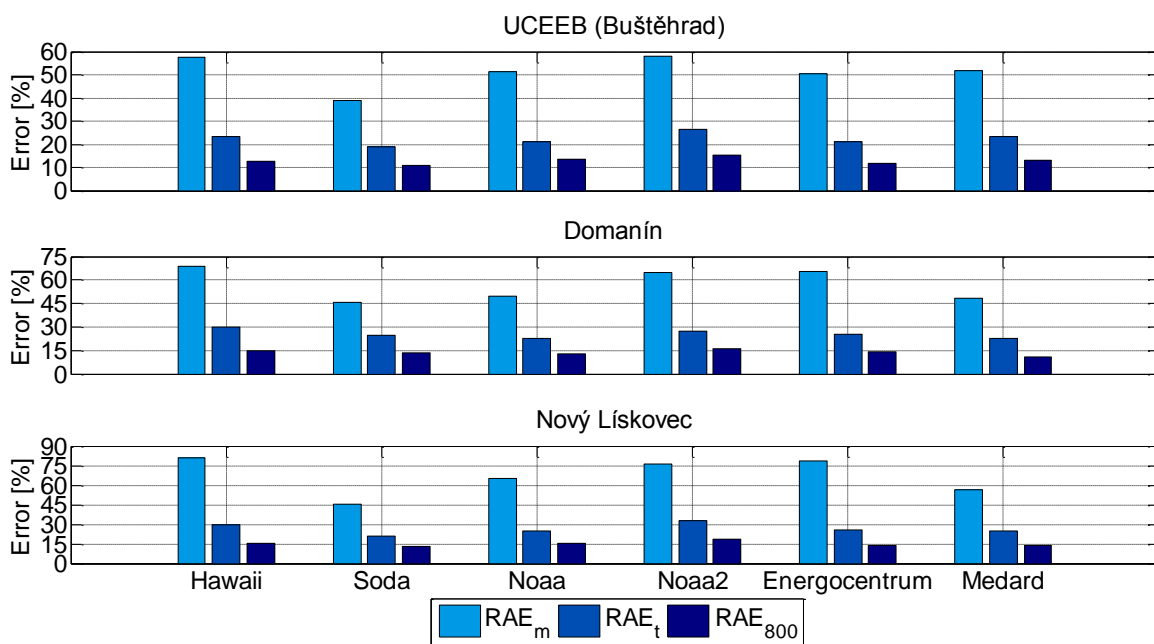


Graf 3: Graf přesnosti předpovědi jednotlivých zdrojů (RAE – průměry za vyhodnocovací období).

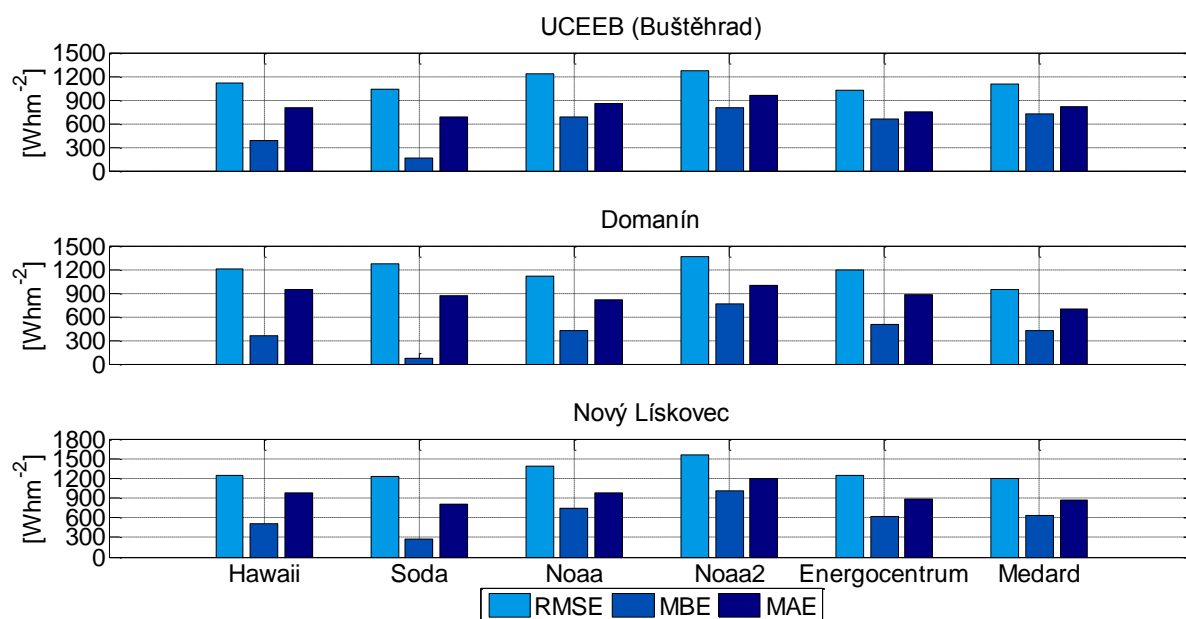


Graf 4: Graf přesnosti předpovědi jednotlivých zdrojů (RMSE, MBE, MAE).

b) Přesnost predikce na 25-48 hodin



Graf 5: Graf přesnosti předpovědi jednotlivých zdrojů (RAE – průměry za vyhodnocovací období).



Graf 6: Graf přesnosti předpovědi jednotlivých zdrojů (RMSE, MBE, MAE).

2.2.7.6 Vytvoření prioritního seznamu

Vyhodnocení přesnosti vedlo k vytvoření prioritního seznamu, který algoritmus využívá k výběru zdroje pro výpočet výsledné predikce. Výsledný seznam:

- Soda
- Hawaii
- NOAA
- NOAA2
- Medard
- Energocentrum

Zdroje Medard a Energocentrum jsou sice přesnější, než zdroje NOAA a NOAA2, ale vzhledem k tomu, že v současné chvíli poskytují pro testovací potřeby algoritmu předpověď pro pouze tři resp. jednu lokalitu v České republice, byly zařazeny na konec seznamu.

Podle tohoto seznamu bylo každému zdroji přiřazeno číslo od 1 do 6 podle pořadí v seznamu. Tyto hodnoty jsou v algoritmu uloženy pod názvem *grades*.

3 SAMOUČÍCÍ ALGORITMUS

Stěžejním výsledkem bakalářské práce je samoučící algoritmus, který zlepšuje dostupnou predikci na základě zpětné vazby v podobě historicky měřených dat. Celý algoritmus je naimplementován ve vývojovém prostředí Matlab [21] firmy MathWorks. Následující kapitola se zabývá popisem tohoto algoritmu.

3.1 Obecný popis

Hlavní funkcí celého algoritmu je funkce *final*. Tato funkce zajišťuje spuštění výpočtu predikce v přednastavený čas, určuje, pro které lokality bude předpověď počítána atd. Před spuštěním této funkce musejí být nastaveny následující parametry:

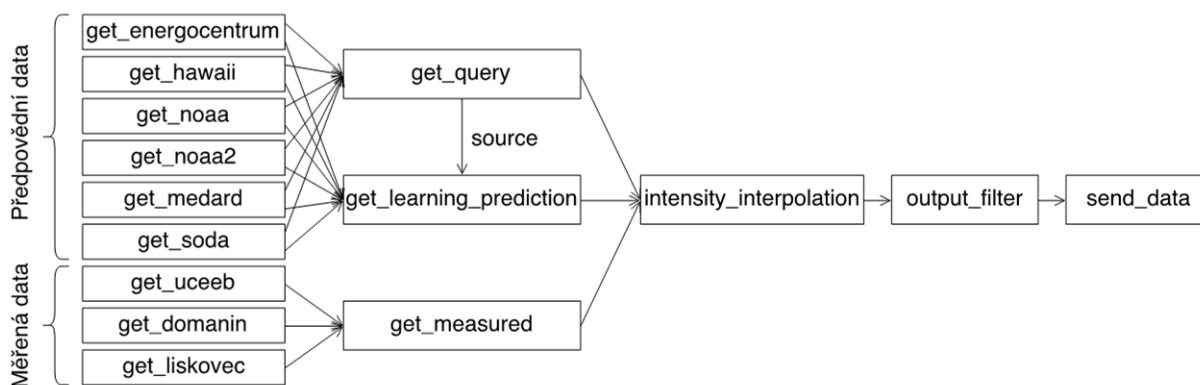
- *learning_period* – určuje, z jak dlouhého období se algoritmus „učí“ [dny]
- *prediction_hours* – určuje, v kterou hodinu/hodiny se spustí výpočet predikce
- *prediction_span* – určuje, na kolik hodin dopředu se bude predikce počítat

Aktuálně používané hodnoty těchto parametrů:

- *learning_period* – 14
- *prediction_hours* – 11 a 23 (predikce se počítá dvakrát denně)
- *prediction_span* – 48

Přestože je predikce počítána dvakrát denně, začátek predikce se nemění a je fixně stanoven na nejbližší budoucí půlnoc.

Funkce *final* dále spouští ostatní funkce a to dle diagramu, který obsahuje Obrázek 1. Jednotlivé kroky algoritmu jsou popsány v následujících kapitolách.



Obrázek 1 Diagram znázorňující průběh samoučícího algoritmu.

3.2 Funkce pro získání předpovědních a měřených dat

První částí algoritmu jsou funkce typu *get_xxx*, kde *xxx* značí daný zdroj predikce či měřených dat. Tyto funkce mají za úkol sjednotit rozdílné vlastnosti jednotlivých zdrojů a vrátit předpovědní resp. měřená data pro den, který je vložen jako vstupní parametr. Dalším vstupním parametrem funkcí týkajících se předpovědních dat jsou GPS souřadnice polohy, pro kterou je předpověď aktuálně počítána. Výstupem těchto funkcí jsou data v jednotném formátu: 24 hodnot (vždy v celou hodinu) a k nim příslušné časové značky. Data jsou v jednotkách W/m^2 . Dalším výstupem těchto funkcí je parametr *data_condition*. Tento parametr může nabývat několika hodnot, pomocí kterých informuje nadřazenou funkci o stavu vrácených dat.

Další úlohou těchto funkcí je ukládání předpovědí pro další použití, aby se předešlo opakovanému stahování stejné předpovědi po několik dní za sebou. Vzhledem k tomu, že stahování předpovědi je časově nejnáročnější část celého algoritmu, je implementace této části velmi důležitá z hlediska efektivity. Data jsou uložena ve speciálním typu proměnné zvané persistentní proměnná [32]. Tato proměnná zajistí, že data zůstanou v proměnné uložena i po ukončení a opětovném otevření funkce a zároveň (na rozdíl od globální proměnné) umožní stejné pojmenování proměnné ve více funkcích. Data v této proměnné zůstávají uložena, dokud jsou potřeba, tedy po dobu určenou parametrem *learning_period*.

3.3 Výběr zdroje předpovědních dat

Výběr nejlepšího dostupného zdroje předpovědních dat na den, pro který je předpověď počítána, zajišťuje funkce `get_query`. Funkce využívá seznam předpovědních zdrojů `grades`, jehož sestavení je popsáno v kapitole 2.2.7 a funkce pro získání předpovědních dat popsané v minulé kapitole.

Funkce `get_query` postupuje podle seznamu `grades` a volá příslušné funkce pro získání předpovědních dat. Po stažení předpovědi z daného zdroje může nastat několik situací:

- V případě, že jsou dle parametru `data_condition` předpovědní data z daného zdroje konzistentní (žádná nechybí), funkce tuto předpověď, společně s informací ze kterého zdroje pochází, posílá k dalšímu zpracování.
- Dále se může stát, že data z daného zdroje nebudou z nějakého důvodu dostupná. Tento stav může nastat například při výpadku na serveru, nebo chybě při stahování. V tomto případě funkce `get_query` postupuje dál v seznamu `grades` a zkouší následující zdroj.
- Poslední možností je, že parametr `data_condition` má hodnotu odpovídající situaci, že ve stažených datech některé informace chyběly a byly nahrazeny nulami. V tomto případě jsou data uložena a algoritmus dále postupně zkouší další zdroje. Jsou-li ostatní zdroje nedostupné, budou tato data použita pro výpočet zpřesněné předpovědi.

Po vybrání nejlepšího možného zdroje predikovaných dat na následující den je nutné získat predikci tohoto zdroje na několik minulých dnů. Počet těchto dnů je nazýván učící období a je určen parametrem `learning_period`. Získání těchto zpětných predikcí je úkolem funkce `get_learning_prediction`.

Tato funkce má opět k dispozici funkce pro získání předpovědních dat a seznam předpovědních zdrojů `grades`. Dále je jí funkcí `get_query` předána informace o použitém zdroji. Funkce primárně stahuje data ze zdroje, který byl použit ve funkci `get_query`. Je-li tento zdroj pro nějaký den nedostupný, je pro daný den použit nejlepší dostupný zdroj. Tato funkce také bere v potaz hodnotu parametru `data_condition` a na jeho hodnotu reaguje obdobně jako funkce `get_query`.

3.4 Výběr zdroje měřených dat

Příslušný zdroj měřených dat je vybírán funkcí `get_measured`. Tato funkce nejprve vypočítá vzdálenost bodu, pro který je předpověď počítána, od všech bodů, kde jsou k dispozici měřená data. Tento výpočet je popsán v kapitole 3.4.1. Funkce tedy pro každý bod vytváří seznam, který je obdobou seznamu `grades` z funkcí `get_query` a `get_learning_prediction`. Poté je automaticky vybrán zdroj, kterému odpovídá nejkratší vypočítaná vzdálenost (pokud je tento k dispozici).

Tato funkce následně využívá další funkce pro získání měřených dat. Chování funkce `get_measured` je parametrem `data_condition` ovlivněno obdobně jako chování funkcí `get_learning_prediction` a `get_query`. Pokud je nejbližší zdroj nedostupný, je použit další v pořadí.

3.4.1 Výpočet vzdálenosti dvou lokalit

Výpočet vzdálenosti dvou bodů na kulové ploše (Země) je ze zjevných důvodů složitější, než výpočet vzdálenosti dvou bodů v rovině. Vzhledem k tomu, že vzdálenosti jsou počítány pouze v rámci České republiky a většinou pouze v rámci její části, byly důsledky zakřivení Země na

vypočtenou vzdálenost zanedbány. Všechny výpočty jsou počítány dle vzorců pro rovnou plochu. Pro výpočty tedy lze využít Pythagorovu větu v její neupravené podobě.

Nelze však jednoduše dosadit rozdíly zeměpisných šířek a délek. Délka oblouku opsaného jedním šířkovým stupněm (s_{lat}) totiž není rovna délce oblouku opsané jedním délkovým stupněm (s_{lon}). Toto zhruba platí pouze v okolí rovníku. Je-li tvar planety Země aproximován koulí, mění se délka oblouku opsaného jedním délkovým stupněm v závislosti na zeměpisné šířce. Délka oblouku opsaného jedním šířkovým stupněm je konstantní. Tuto hodnotu je možno spočítat dle vzorce:

$$s_{lat} = \frac{\pi}{180} R,$$

kde:

- s_{lat} – délka oblouku opsaného jedním šířkovým stupněm [km]
- R – průměrný poloměr Země (6371 km)

Po dosazení vychází hodnota s_{lat} rovna 111,194 km. Jak již bylo řečeno, tato hodnota je konstantní pro všechny hodnoty zeměpisné šířky, je tedy použita v algoritmu pro přepočtení rozdílu zeměpisných šířek dvou bodů na hodnotu v kilometrech.

Délka oblouku opsaného jedním délkovým stupněm se mění v závislosti na zeměpisné šířce a to dle následujícího vzorce:

$$s_{lon} = \frac{\pi}{180} R \cos \varphi,$$

kde:

- s_{lon} – délka oblouku opsaného jedním délkovým stupněm [km]
- R – průměrný poloměr Země (6371 km)
- φ – zeměpisná šířka [°]

Vzhledem k tomu, že hodnota s_{lon} je proměnná v závislosti na zeměpisné šířce, je tento vzorec pro přepočet mezi rozdílem zeměpisných šířek dvou bodů a vzdáleností v kilometrech přímo implementován ve funkci `get_measured`. [22][23]

3.5 Úprava měřených dat

Tato kapitola se zabývá úpravou měřených dat. Data je nutné upravit z několika důvodů. Jeden z nich byl nastíněn již v kapitole 2.1. Tímto důvodem je fakt, že se v datech vyskytují špičky, které výrazně převyšují průměr blízkých hodnot. Řešením je nekauzální klouzavý průměr, jehož výpočet je popsán v následující kapitole.

3.5.1 Výpočet nekauzálního klouzavého průměru

Nekauzální klouzavý průměr, který je pouhou úpravou kauzálního klouzavého průměru [5], se vypočítá dle následujícího vzorce:

$$\bar{x}_i = \frac{1}{n} \sum_{k=i-\frac{n}{2}}^{i+\frac{n}{2}} x_k$$

$$i \in \left\langle \frac{n}{2}; m - \frac{n}{2} \right\rangle$$

kde:

- \bar{x}_i – i-tá hodnota ve výsledné posloupnosti dat (průměry)
- n – počet hodnot, ze kterých je průměr počítán
- x_k – k-tá hodnota v původní posloupnosti dat
- m – počet původních hodnot (v tomto případě 1440)

Hodnota n byla zvolena 60. Průměr je tedy vždy počítán z dat naměřených během půl hodiny před a půl hodiny po daném čase. Ze vzorců je vidět, že ve výsledných datech chybí $\frac{n}{2}$ hodnot na začátku a $\frac{n}{2}$ hodnot na konci oproti původním datům. Vzhledem k tomu, že tyto „začátky“ a „konce“ připadají na časy okolo půlnoci, je možné tyto chybějící hodnoty nahradit nulami.

3.5.2 Kompenzace nízkého počtu zdrojů měřených dat

Ideální by bylo mít měřená data z každého bodu, pro který je počítána předpověď. V současné době jsou však k dispozici data z pouhých tří poloh. Tento problém byl částečně vyřešen následujícím způsobem.

Po výběru zdroje měřených dat dle způsobu popsaného v kapitole 3.4, jsou spočítány rozdíly za celé učící období mezi měřenými daty a daty predikovanými pro místo měření. Celkem je tedy získáno: délka učícího období \times 24 hodnot. Tyto rozdíly jsou poté přičteny k predikovaným hodnotám pro polohu, pro kterou je aktuálně počítána předpověď. Takto upravená predikovaná data jsou použita jako zpětná vazba místo chybějících dat měřených pro danou polohu.

Účinnost této metody byla potvrzena následujícím experimentem. Pro období od 10.1.2016 do 26.4.2016 (107 dní) byly počítány předpovědi pro polohy Domanín (50° N, 14,7° E) a Nový Lískovec (49,2° N, 16,5° E) celkem třemi způsoby. První způsob používá pro zpětnou vazbu neupravená měřená data z meteorologické stanice UCEEB. Druhý způsob používá pro zpětnou vazbu data z UCEEBu upravená pomocí metody popsané v předchozím odstavci. Třetí způsob je pouze referenční a používá data naměřená stanicemi v Domaníně (resp. v Novém Lískovci).

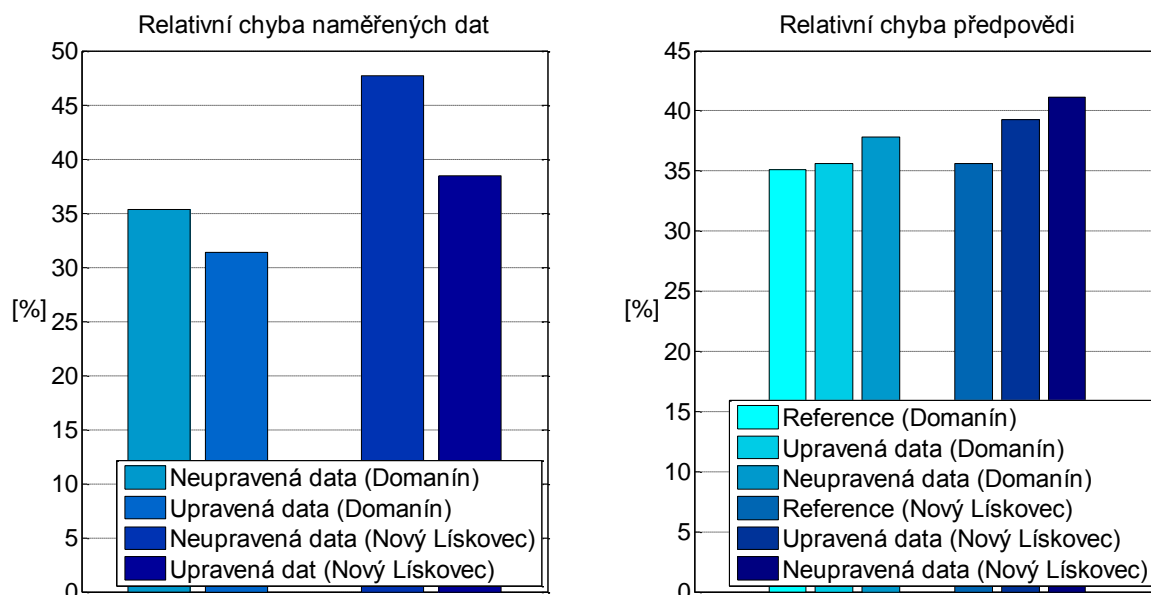
Výsledky tohoto pokusu zobrazuje Graf 7. Levý podgraf obsahuje průměrné relativní chyby upravených a neupravených dat z UCEEBu v jednotlivých lokalitách (Domanín a Nový Lískovec). Pravý podgraf obsahuje průměrné relativní chyby předpovědí počítaných podle způsobů popsaných v minulém odstavci. Všechny relativní chyby byly počítány dle následujícího vzorce:

$$r_m = \frac{|s_p - s_m|}{s_m} \cdot 100$$

kde:

- s_p – denní suma predikovaných hodnot v daný den [Wh/m²],
- s_m – denní suma naměřených hodnot v daný den [Wh/m²],
- r_m – relativní chyba vztažená k denní sumě naměřených hodnot v daný den [%].

Tímto způsobem bylo pro každý způsob předpovědi a každá upravená a neupravená naměřená data získáno celkem 107 hodnot, které odpovídají 107 dnům, na kterých byla přesnost testována. Z těchto hodnot byl poté vypočítán průměr. Z grafů je evidentně vidět, že přesnost předpovědi se po zavedení této kompenzace zvýšila.



Graf 7: Relativní chyby upravených a neupravených naměřených dat z UCEEBu a výsledné předpovědi v závislosti na použitých naměřených datech v jednotlivých lokalitách.

3.5.3 Korekce nenulových hodnot v nočních hodinách

Měřená data v některých případech obsahují nízké nenulové hodnoty i v nočních hodinách. Vzhledem ke zvolenému způsobu výpočtu výsledné předpovědi je nutné tyto hodnoty odstranit, protože by se vyskytly i ve výsledné předpovědi.

Největším problémem při odstraňování těchto hodnot je správně určit časové období, ve kterém by se v datech nenulové hodnoty neměly vyskytovat. Toto období je totiž v průběhu roku pro každý den jiné a je dáno východem a západem Slunce. Tento problém byl vyřešen použitím funkce *suncycle* [24], která pro zadanou polohu GPS a den v roce vrací mimo jiné právě čas východu a západu Slunce.

Nyní již stačí všechny hodnoty v daném dni před tímto časem (resp. po tomto čase) nastavit na nulu.

3.6 Úprava předpovědních dat

Jak již bylo nastíněno v kapitole 2.2, hlavním nedostatkem většiny zdrojů předpovědních dat je posunutí a „roztažení“ předpovědi tak, že se nulové hodnoty intenzity slunečního záření v predikci objevují až v cca 23 hodin. Tento problém se vyskytuje u zdrojů Soda, Energo centrum, Hawaii a NOAA. Pokud by data nebyla upravena ještě před vstupem do hlavní části algoritmu, výrazně by ovlivnila výslednou předpověď. Řešení problému je následující:

Každé hodnotě předpovědi je přiřazena požadovaná hodina, která má dané hodnotě odpovídat. Tyto hodnoty jsou pro každý měsíc jiné. Odpovídá-li například nejvyšší hodnota původní neupravené únorové předpovědi sedmnácté hodině, je této hodnotě nově přiřazena 13. hodina. Zapadá-li Slunce podle původní předpovědi v 23 hodin, upraví se tato hodnota na 16. hodinu. Tento postup samozřejmě nezajistí úplnou přesnost předpovědi, ale odstraní převážnou část chyby způsobené posunem a „roztažením“ předpovědi.

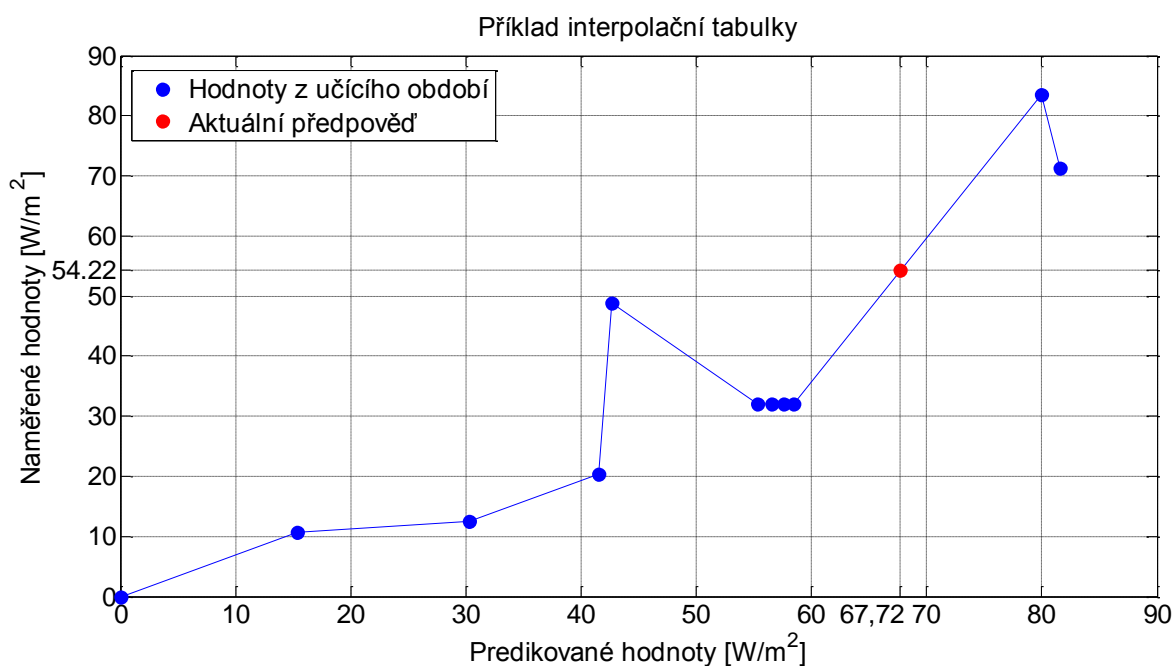
Vzhledem k tomu, že většina předpovědí má tříhodinový interval mezi daty, je ze všech předpovědí (upravených, i úpravu nepotřebujících) lineární interpolací získáno 24 hodnot na

jeden den s hodinovými intervaly. Tímto je formát předpovědních dat sjednocen a data mohou v algoritmu postoupit dále.

3.7 Interpolace

Samotný výpočet výsledné předpovědi zajišťuje funkce *intensity_interpolation*. Výpočet je založen na lineární interpolaci. Naměřená i predikovaná data z učicího období jsou rozdělena podle hodiny, pro kterou byla predikována (resp. pro kterou byla naměřena). Vzniká tedy celkem dvakrát (naměřené a predikované hodnoty) 24 skupin s počtem hodnot odpovídajícím délce učicího období. V případě, že se v jedné skupině predikovaných hodnot vyskytuje jedna hodnota více než jednou, jsou tyto duplicitní výskyty odstraněny a dané hodnotě je přiřazena průměrná hodnota z naměřených hodnot odpovídajících duplicitním hodnotám. Tyto skupiny jsou dále uspořádány do 24 tabulek, kde na jedné ose jsou hodnoty naměřené a na druhé ose hodnoty predikované. Do každé tabulky je navíc přidáván bod (0,0), který zajistí vypočtení hodnoty i pro případ, že by nově predikovaná hodnota byla nižší, než všechny predikované hodnoty z učicího období. V případě, že je nově predikovaná hodnota vyšší než všechny hodnoty predikované během učicího období, je vrácena nejvyšší hodnota během tohoto období naměřená. Pomocí těchto 24 interpolačních tabulek a 48 hodnot získaných ze zdroje předpovědních dat, je lineární interpolací získáno 48 výsledných hodnot. [35]

Graf 8 obsahuje příklad jedné z těchto interpolačních tabulek, konkrétně tabulky použité pro výpočet hodnoty předpovědi pro 9. hodinu 12.1.2016. Na grafu je červenou barvou zvýrazněna hodnota předpovídáná jedním ze zdrojů predikčních dat (osa X) – $67,72 \text{ W/m}^2$ a výsledná vypočítaná předpověď (osa Y) – $54,22 \text{ W/m}^2$.



Graf 8: Příklad interpolační tabulky (výpočet předpovědi pro 12.1.2016 hodina 9).

3.8 Výstupní filtr

Výsledná předpověď je před odesláním na server kontrolována funkcí *output_filter*. Výstupní filtr byl implementován z toho důvodu, aby nemohl nastat případ, kdy se do výsledné předpovědi nějakou chybou (předpovědních, či měřených dat) dostane nesmyslně nízká či vysoká předpověď.

Tato funkce porovnává denní sumu dnů, na které je předpovídáno s maximální a minimální denní sumou z hodnot naměřených během posledních 30 dní. Funkce *output_filter* obsahuje parametry *max_ratio* a *min_ratio*, které označují maximální (resp. minimální) povolený poměr mezi denní sumou aktuálně predikovaných hodnot a maximální (resp. minimální) denní sumou z hodnot naměřených během posledních 30 dní. Je-li poměr vyšší (resp. nižší), než tyto parametry, je předpověď nahrazena denním průběhem s maximální (resp. minimální) sumou naměřenou během posledních 30 dní.

Aktuální hodnoty parametrů jsou:

- *max_ratio* – 1,65
- *min_ratio* – 0,4

Hodnoty parametrů byly nastaveny tak, aby nedošlo k nahrazení správné předpovědi, ale aby byly vyfiltrovány a nahrazeny vadné (příliš vysoké, či příliš nízké) předpovědi.

3.9 Odesílání dat

Data jsou odeslána na server až po kompletním vypočítání předpovědi pro všechny body mřížky. Odesílání dat zajišťuje funkce *send_data*. Tato funkce postupně odesílá data pro jednotlivé body mřížky. Po odeslání dat pro jeden bod čeká na odpověď serveru. Je-li odpověď kladná („OK“), odesílá předpověď pro další bod. Vyskytne-li se chyba (odpověď serveru je „FAIL“), funkce zkouší tato data pětkrát za sebou odeslat ve čtvrt-vteřinových intervalech. Nepodaří-li se data během těchto pěti pokusů odeslat, je pokus opakován každou minutu, dokud nenastane čas výpočtu další předpovědi.

Společně s daty je odeslán speciální kód vytvořený pomocí funkce *DataHash* [25], využívající datum, na které je předpověď počítána. Jedná se o jednoduchý způsob zamezující nahrání dat na server neoprávněnými osobami.

3.10 Ochrana proti výpadkům

Není-li k dispozici některý ze zdrojů předpovědních dat, řeší to funkce *get_query* způsobem popsaným v kapitole 3.3. Pokud však není k dispozici žádný ze zdrojů, vrací funkce *get_query* předpověď uloženou při posledním bezproblémovém získávání předpovědi.

Problém výpadku dat se však netýká pouze předpovědních dat, ale i dat měřených. Vzhledem ke způsobu výpočtu předpovědi je nutné, aby počet měřených dat získaných během učícího období byl stejný jako počet predikčních dat z učícího období. Vznikne-li výpadek v měřených datech, musí se příslušný den odstranit i z predikčních dat a naopak. Poté je zajištěno, že celý algoritmus bude pracovat správně.

Při výběru predikčních a měřených dat je také kontrolováno, zda data neobsahují nesmyslné hodnoty. Pokud se v datech vyskytne mezi 10. a 15. hodinou hodnota nižší, než 10 W/m^2 , jsou data označena za špatná a data pro daný den jsou získána z následujícího nejlepšího zdroje.

4 VÝSLEDKY SAMOUČÍCÍHO ALGORITMU

4.1 Dosažená přesnost

Přesnost výsledné predikce byla vyhodnocena stejnými způsoby jako přesnost jednotlivých zdrojů predikce v kapitole 2.2.7. Dosažená přesnost a porovnání s ostatními zdroji je viditelná na grafech v následujících kapitolách. Vyhodnocení přesnosti proběhlo až po úpravě dat

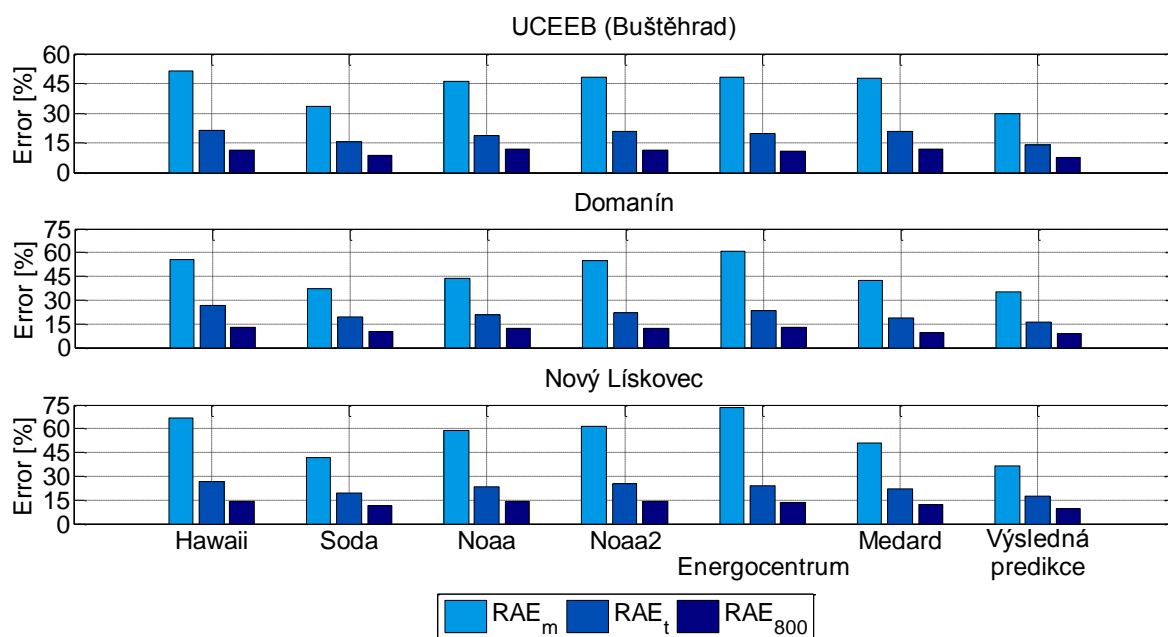
způsoby popsány v kapitole 3.6. Přesnost původních, neupravených dat by byla vzhledem k posunu a „roztážení“ mnohem nižší.

Z grafů je patrné, že výsledná predikce samoučícího algoritmu je v drtivé většině případů mnohem přesnější, než zdroje původních predikovaných hodnot. Jedinou výjimku tvoří předpověď na 25-48 hodin, kde byla v oblasti Domanín hodnota MBE výsledné predikce o cca 30 Wh/m² vyšší, než hodnota MBE zdroje Soda.

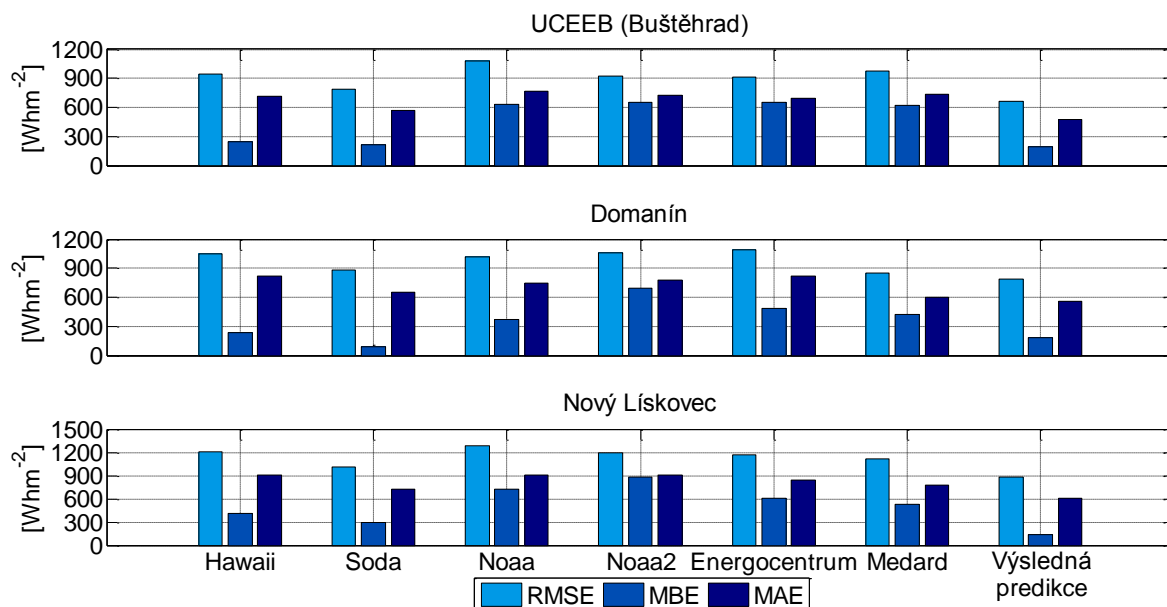
Předpověď pro		UCEEB (Buštěhrad)		Domanín		Nový Lískovec	
Počet hodin dopředu		0-24	25-48	0-24	25-48	0-24	25-48
průměr	RAE_m [%]	29,8	37,2	35,1	42	36,7	46
	RAE_t [%]	13,7	16,8	16,2	21,1	17,3	19,6
	RAE_{800} [%]	7,46	9,51	8,8	11,6	9,53	11,4
medián	RAE_m [%]	19,3	23,8	21,8	29,1	19,4	27,3
	RAE_t [%]	10,9	13,6	12,6	16,1	11,5	15,7
	RAE_{800} [%]	5,02	6,32	5,95	9,42	6,27	7,58
$RMSE$ [Wh/m ²]		663	853	790	985	876	1020
MBE [Wh/m ²]		194	53,2	186	108	138	184
MAE [Wh/m ²]		471	601	556	735	602	718

Tabulka 3: Přesnost výsledné predikce podle jednotlivých způsobů vyhodnocení, horizontu predikce a polohy.

4.1.1 Přesnost predikce na 0-24 hodin

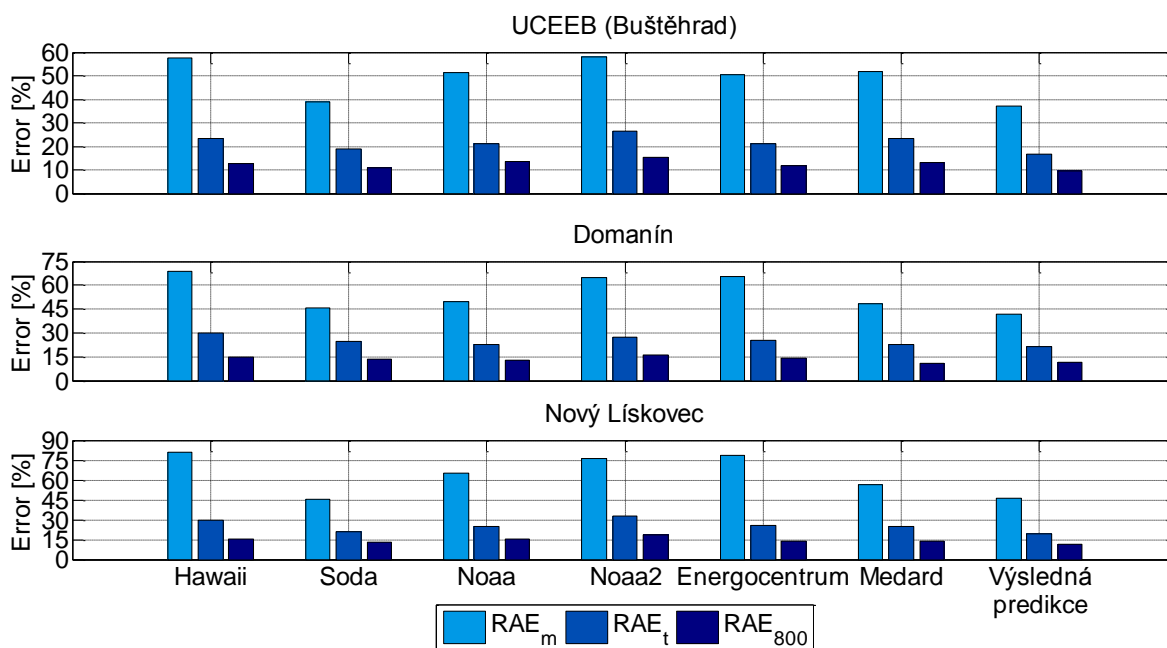


Graf 9: Graf přesnosti předpovědi jednotlivých zdrojů (RAE – průměry za vyhodnocovací období).

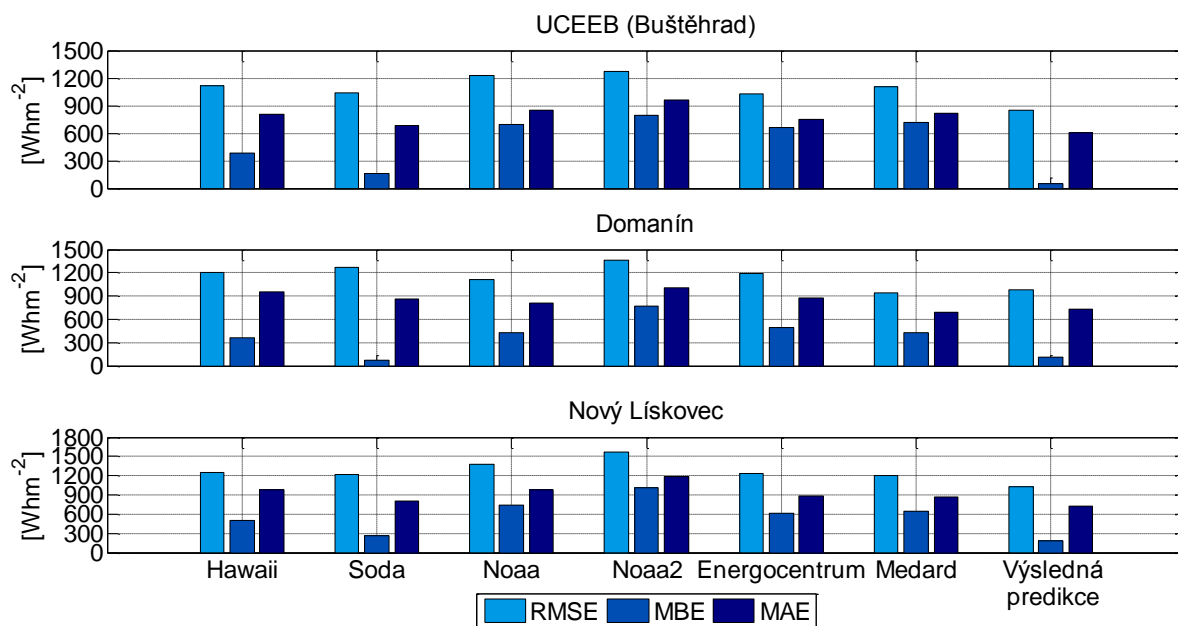


Graf 10: Graf přesnosti předpovědi jednotlivých zdrojů (RMSE, MBE, MAE).

4.1.2 Přesnost predikce na 25-48 hodin



Graf 11: Graf přesnosti předpovědi jednotlivých zdrojů (RAE – průměry za vyhodnocovací období).



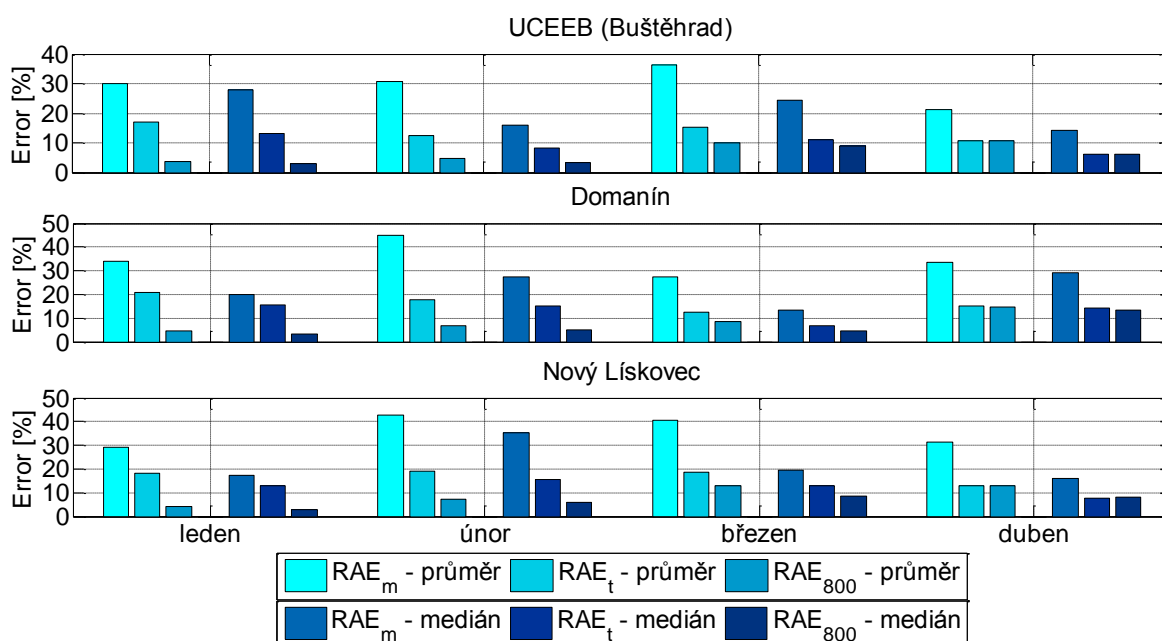
Graf 12: Graf přesnosti předpovědi jednotlivých zdrojů (RMSE, MBE, MAE).

4.1.3 Přesnost výsledné predikce po měsících

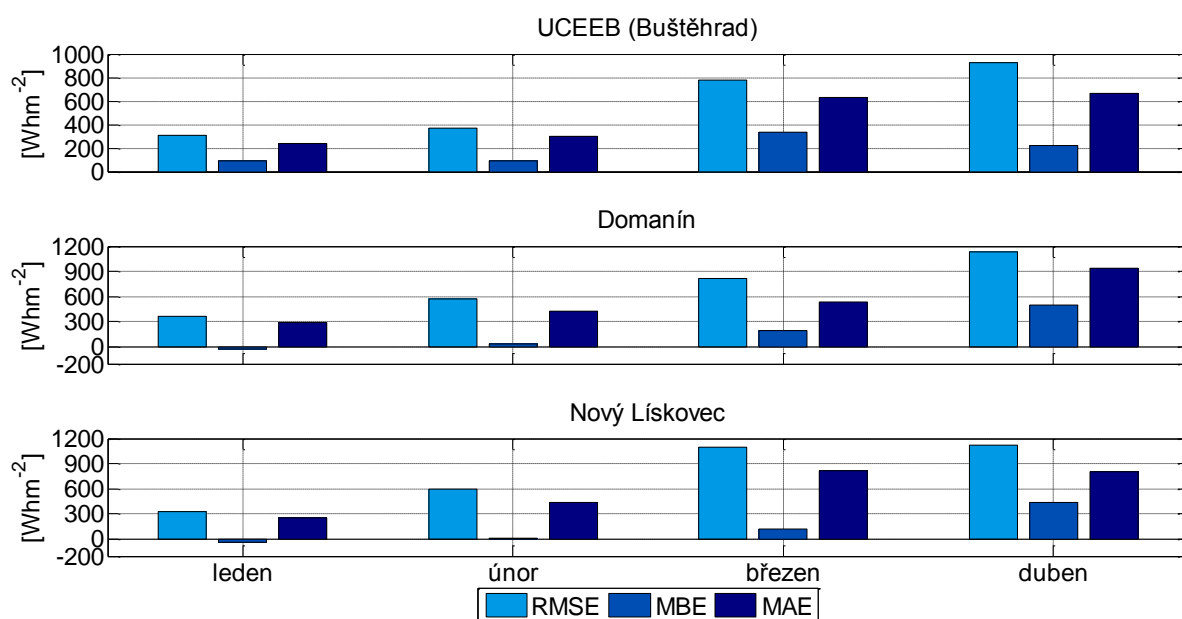
Přesnost výsledné predikce byla vyhodnocena také v rámci jednotlivých měsíců. Toto zobrazení pomůže zjistit vývoj chyby v průběhu roku. Z následujících grafů (převážně z grafů Graf 14 a Graf 16) je patrné, že predikce byla nejpřesnější v měsíci lednu a její přesnost v následujících měsících postupně klesá.

Vyhodnocení za měsíc leden obsahuje pouze data z 22 dní (10.1. – 31.1.) a měsíc duben byl vyhodnocen pomocí dat z 27 dní (1.4. – 27.4.).

4.1.3.1 Predikce na 0-24 hodin

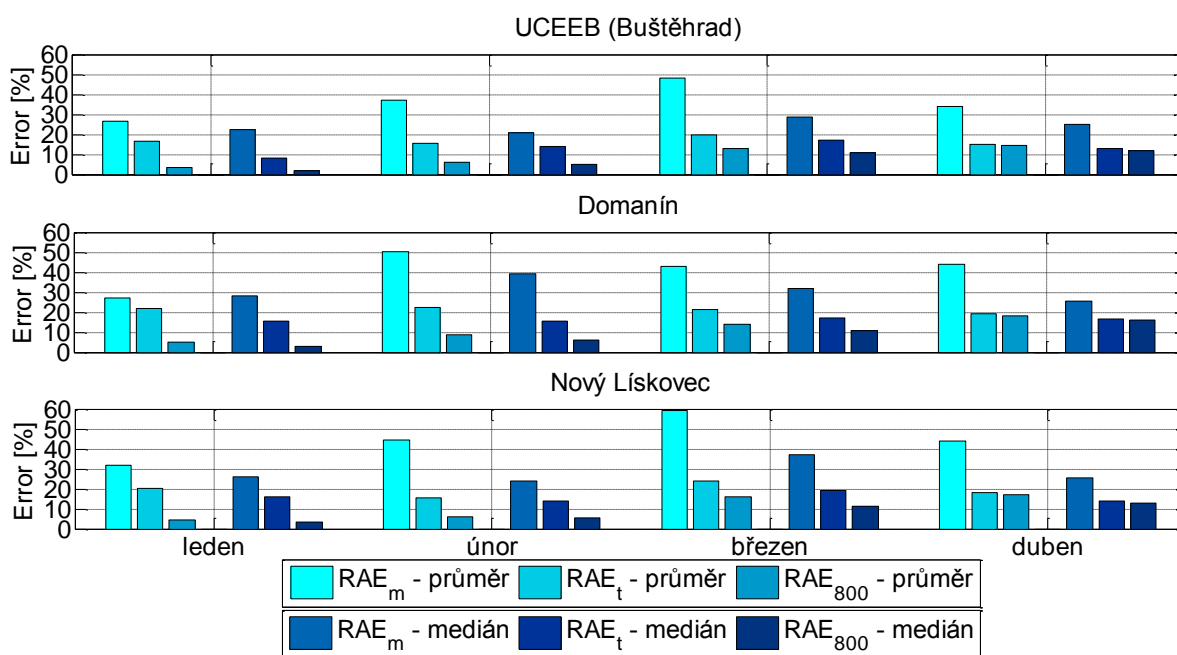


Graf 13: Graf přesnosti výsledné předpovědi (RAE – průměry a mediány za jednotlivé měsíce).

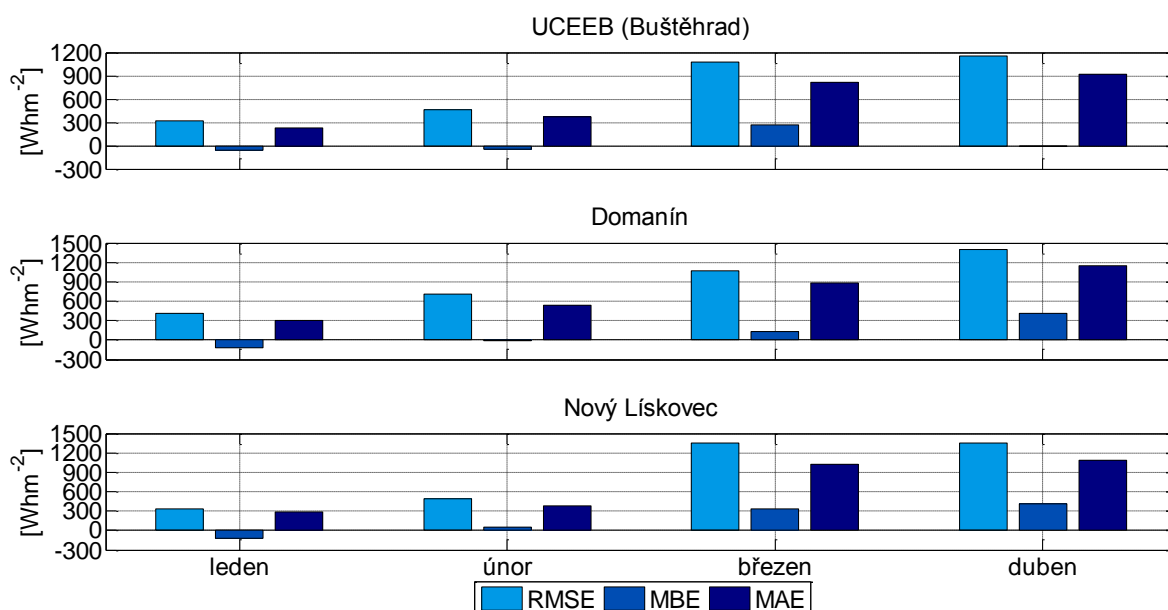


Graf 14: Graf přesnosti výsledné předpovědi (RMSE, MBE, MAE za jednotlivé měsíce).

4.1.3.2 Predikce na 25-48 hodin



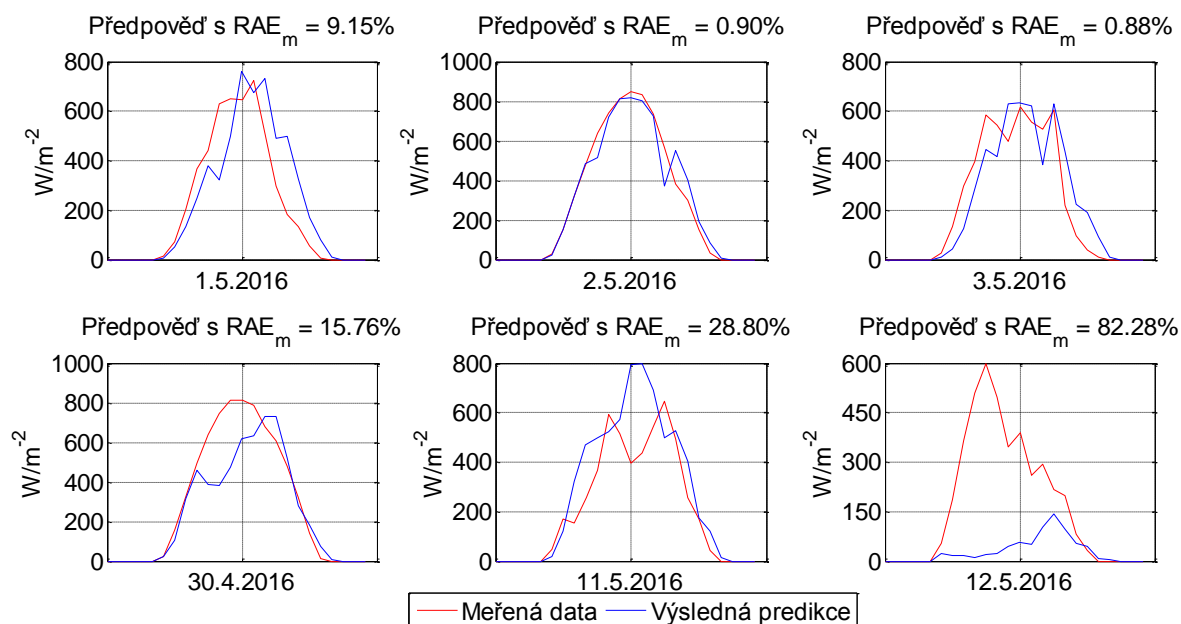
Graf 15: Graf přesnosti výsledné předpovědi (RAE – průměry a mediány za jednotlivé měsíce).



Graf 16: Graf přesnosti výsledné předpovědi (RMSE, MBE, MAE za jednotlivé měsíce).

4.2 Příklady predikce

Následující graf ukazuje příklady výsledných předpovědí. Vzhledem k tomu, že RAE_m bylo počítáno pro denní sumy, je přesnost předpovědi na 2.5. a 3.5. srovnatelná i přes evidentní posun v předpovědi pro 3.5.



Graf 17: Příklady předpovědi pro polohu $50^\circ N$, $14^\circ E$ (cca UCEEB).

5 DOSTUPNOST PŘEDPOVĚDI

Celý algoritmus byl implementován jako služba pod názvem PV Forecast [33]. Výpočet předpovědi je spuštěn na serveru v Univerzitním centru energeticky efektivních budov ČVUT (UCEEB) v Buštěhradě. Z tohoto serveru jsou vypočtené předpovědi posílány na jiný server na UCEEBu. Posílání je popsáno v kapitole 3.9. Tento server data přijme a uloží.

Služba PV Forecast je placená, přístup k datům je tedy zákazníkům povolen po zadání unikátního kódu. Vzhledem k tomu, že se jedná o službu zaměřenou na strojové využívání předpovědi, jsou data k dispozici pouze pomocí HTTP dotazu. Dotaz vracející předpověď má následující formát:

```
http://www.pvforecast.cz/api/?key=hc8217&lat=50.157&lon=14.170
&format=simple&type=hour&number=24&date=1
```

Význam jednotlivých parametrů je popsán v následující tabulce (Tabulka 4).

Klíč	Význam	Možné hodnoty	Výchozí hodnota
key*	Unikátní uživatelský klíč	Šestimístný kód přidělený při registraci	
lat*	Zeměpisná šířka	Číslo vyplněné při registraci (tři desetinná místa)	
lon*	Zeměpisná délka	Číslo vyplněné při registraci (tři desetinná místa)	
format	Formát dat	simple/csv/json	simple
type	Hodinové nebo denní sumy	hour/day	hour
number	Délka předpovědi v hodinách nebo dnech (počet vácených hodnot) **	type = hour: 24/48/72 *** type = day: 1/2/3 ***	24 (type = hour) 1 (type = day)
date	Umožňuje vypnout časovou značku u formátu Simple. U formátů JSON a CSV nemá význam.	1 = zapnuto 0 = vypnuto	1
dst	Automatická volba letního času (daylight saving time)	1 = zapnuto 0 = vypnuto	1
<p>* Povinné parametry.</p> <p>** Pokud předpověď na danou dobu není dostupná, je v odpovědi místo příslušné hodnoty prázdné místo (formáty Simple a CSV) nebo hodnota NULL (formát JSON).</p> <p>*** Předpověď na 72 hodin (3 dny) není podporována a jedná se o přípravu na možné budoucí rozšíření služby.</p>			

Tabulka 4: Tabulka popisující význam jednotlivých parametrů dotazu pro přístup k předpovědi.

Data je možné odeslat ve formátu csv, json nebo simple. Formát simple obsahuje pouze jedinou časovou značku (začátek předpovědi). Poté následují data oddělená svíslou čarou. Příklad výstupu:

```
2016-04-12|0|0|0|0|0|0|0|15|80|262|512|538|507|481|565|532|235
|313|176|108|28|0|0|0
```

Zákazník má omezený počet dotazů na jednu předpověď (konkrétně 5). Jedinou výjimkou je prvních 14 dní od vytvoření nového uživatelského účtu, kdy počet dotazů není omezen.

V závislosti na času dotazu jsou vráceny různé předpovědi a to dle následující tabulky (Tabulka 5). Je-li dotaz odeslán v době mezi 12:00 a 0:00, vrací server předpověď vypočtenou v 11:00. V případě dotazu mezi 0:00 a 12:00, vrací server předpověď vypočtenou v 23:00. Začátek a rozsah předpovědi ale zůstává stejný.

12:00	18:00	0:00	6:00	12:00	18:00	0:00	6:00	12:00	18:00	0:00
Dotaz v této době										
12:00 - 0:00 a 0:00 - 12:00										
				Rozsah předpovědi						

Tabulka 5: Tabulka popisující začátek a rozsah předpovědi v závislosti na času dotazu.

Aktuální předpověď je k nahlédnutí na webové adrese <http://www.pvforecast.cz/predpoved-body.php>. [26]

6 MOŽNOSTI KE ZLEPŠENÍ A BUDOUCÍ PRÁCE

- Seznam předpovědních zdrojů *grades*, zmíněný v kapitole 2.2.7.6, je v současnou chvíli fixní. Jednou z plánovaných úprav algoritmu je vytvořit adaptivní funkci, která bude vyhodnocovat přesnost zdrojů v reálném čase a na základě jejich výsledků se bude tento seznam upravovat. Výsledkem by pak mohlo být například střídání různých seznamů podle ročního období či měsíců, pokud by se v budoucnu ukázalo, že některé zdroje jsou v různých měsících přesnější než jiné.
- Vzhledem k tomu, že celý výpočet je založen na zpětné vazbě tvořené historicky měřenými daty, je nutné mít k dispozici co nejhustší síť meteorologických stanic. V současné době využívá algoritmus data z pouhých tří stanic, která jsou pro použití v jiných polohách v rámci České republiky upravována způsobem popsáným v kapitole 3.5.2. Do budoucna je tedy nutné navýšit počet měřených bodů v České republice.
- Národní úřad pro oceán a atmosféru (NOAA) používá několik různých modelů pro předpověď stavu atmosféry na celé zeměkouli. Algoritmus v současnou chvíli využívá data pouze ze dvou těchto modelů (NOAA a NOAA2). Využitím dat z více modelů by se snížila pravděpodobnost výpadku všech zdrojů předpovědi najednou. Je také možné, že některý z těchto modelů bude přesnější, než modely dosud používané.

7 ZÁVĚR

Hlavním cílem této práce bylo navrhnout a implementovat samoučící algoritmus, využívající měřená data jako zpětnou vazbu pro úpravu predikce intenzity krátkovlnné radiace. Součástí práce také bylo provést rešerši o dostupných zdrojích predikčních a měřených dat.

V první části práce byla vyhodnocena přesnost jednotlivých zdrojů předpovědních dat pomocí několika metrik. Na základě tohoto vyhodnocení byl vytvořen prioritní seznam, který algoritmus využívá pro výběr nejlepšího zdroje predikčních dat pro výpočet výsledné predikce.

Práce dále obsahuje popis samotného algoritmu založeného na lineární interpolaci. Jednou z hlavních výhod tohoto přístupu je jeho univerzálnost. Vzhledem k tomu, že výsledná předpověď není přímo vypočítávaná z původní predikce, ale interpolována pomocí interpolační tabulky, je možné výslednou predikci získat i z jiných meteorologických veličin. Jedinou podmínkou je vzájemná korelace mezi danou veličinou a intenzitou krátkovlnné radiace. Příkladem takovéto veličiny je oblačnost (neboli cloud cover).

Součástí práce je i vyhodnocení přesnosti výsledné predikce a její porovnání s ostatními zdroji. Z grafů uvedených v této části práce je patrné, že výsledná predikce je v drtivé většině případů mnohem přesnější, než zdroje původních predikovaných hodnot. Jedinou výjimku tvoří předpověď na 25-48 hodin, kde byla v oblasti Domanín hodnota MBE výsledné predikce o cca 30 Wh/m² vyšší, než hodnota MBE zdroje Soda.

Výsledný algoritmus byl implementován jako placená služba pod názvem PV Forecast [33]. Provozovatelem této služby je Univerzitní centrum energeticky efektivních budov ČVUT (UCEEB).

Závěry této práce byly publikovány v článcích „Návrh algoritmů pro získání předpovědi dopadajícího slunečního záření“ [34] a „Využití předpovědi intenzity ozáření pro řízení hybridních fotovoltaických systémů“ [35] a prezentovány na konferencích „Nekonvenční zdroje elektrické energie“ ve Vyškově u Brna a „Alternativní zdroje energie 2016“ v Kroměříži.

Seznam použité literatury

- [1] Šmídek, J.: *Prediktivní řízení malých hybridních solárních systémů*. Diplomová práce ČVUT FEL, květen 2015
- [2] Staněk, K.: *Fotovoltaika pro budovy*, ISBN 978-80-247-4278-6, Grada 2013
- [3] Majer, V. a Hejtmánková, P.: *Získání dat pro krátkodobou predikci výroby elektřiny z fotovoltaických elektráren*, 11th International Conference ENERGY – ECOLOGY – ECONOMY 2012, Tatranské Matliare, květen 2012
- [4] *Produktový list týmu RPI – Univerzitní centrum energeticky efektivních budov ČVUT v Praze* [Online] Dostupné z: http://www.uceb.cz/system/files/souboryredakce/rp1_architektura_a_interakce_fina1.pdf [Přístup získán 15. 5. 2016].
- [5] *Klouzavý průměr* [Online] Dostupné z: <http://mathworld.wolfram.com/MovingAverage.html> [Přístup získán 15. 5. 2016].
- [6] *Webový prohlížeč měřených dat – FIEDLER AMS s.r.o.* [Online] Dostupné z: <https://stanice.fiedler-magr.cz/> [Přístup získán 15. 5. 2016].
- [7] *Meteorologické stanice společnosti FIEDLER AMS s.r.o.* [Online] Dostupné z: <http://www.fiedler.company/cs/aplikace/meteorologicke-stanice/relativni-vlhkost-teplota-vzduchu-destove-srazky-globalni-radiace-tl> [Přístup získán 15. 5. 2016].
- [8] *About grib file format* [Online] Dostupné z: <http://www.wmo.int/pages/prog/www/DPS/FM92-GRIB2-11-2003.pdf> [Přístup získán 21. 5. 2016].
- [9] *About NOAA* [Online] Dostupné z: <http://www.noaa.gov/about-our-agency> [Přístup získán 15. 5. 2016].
- [10] *Global Climate and Weather Modeling Branch: Mission Statement* [Online] Dostupné z: <http://www.emc.ncep.noaa.gov/gmb/mission.html> [Přístup získán 15. 5. 2016].
- [11] *Global Forecast System (GFS) documentation* [Online] Dostupné z: http://nomads.ncep.noaa.gov/txt_descriptions/GFS_doc.shtml [Přístup získán 15. 5. 2016].
- [12] *Informace o modelu „NOAA“* [Online] Dostupné z: http://nomads.ncep.noaa.gov:9090/dods/fnl/fnl20160409/fnlflx_00z.info [Přístup získán 15. 5. 2016].
Poznámka: do odkazu je nutné vložit aktuální datum
- [13] *Informace o modelu „NOAA 2“* [Online] Dostupné z: http://nomads.ncep.noaa.gov:9090/dods/gfs_0p25/gfs20160331/gfs_0p25_00z.info [Přístup získán 15. 5. 2016].
Poznámka: do odkazu je nutné vložit aktuální datum
- [14] *Informace o zdroji Soda* [Online] Dostupné z: http://www.soda-is.com/eng/about/soda_project_factsheet.html [Přístup získán 15. 5. 2016].

- [15] *Služby poskytované zdrojem Soda* [Online] Dostupné z: <http://www.soda-is.com/eng/services/index.html> [Přístup získán 15. 5. 2016].
- [16] *Profil společnosti ENERGOCENTRUM PLUS, s.r.o.* [Online] Dostupné z: <http://www.energocentrum.cz/profil-spolecnosti/> [Přístup získán 15. 5. 2016].
- [17] *Předpověď počasí společnosti ENERGOCENTRUM PLUS, s.r.o.* [Online] Dostupné z: <http://rcware.eu/weather/prague/> [Přístup získán 15. 5. 2016].
- [18] *Nápověda služby Medard* [Online] Dostupné z: <http://www.medard-online.cz/help> [Přístup získán 15. 5. 2016].
- [19] *About PacIOOS* [Online] Dostupné z: <http://oos.soest.hawaii.edu/pacioos/about/> [Přístup získán 15. 5. 2016].
- [20] *Data Access Form - Hawaii* [Online] Dostupné z: http://oos.soest.hawaii.edu/erddap/griddap/NCEP_Global_Best.html [Přístup získán 15. 5. 2016].
- [21] *Vývojové prostředí Matlab* [Online] Dostupné z: <http://www.mathworks.com/products/matlab/> [Přístup získán 15. 5. 2016].
- [22] Osborne, Peter (2013). *Chapter 5: The geometry of the ellipsoid* [Online] Dostupné z: <http://www.mercator99.webspace.virginmedia.com/mercator.pdf> [Přístup získán 15. 5. 2016].
- [23] Rapp, Richard H. (April 1991). *Chapter 3: Properties of the Ellipsoid* [Online] Dostupné z: http://kb.osu.edu/dspace/bitstream/handle/1811/24333/Rapp_Geom_Geod_Vol_I.pdf?sequence=1 [Přístup získán 15. 5. 2016].
- [24] *Informace o funkci suncycle* [Online] Dostupné z: <http://mooring.ucsd.edu/software/matlab/doc/toolbox/geo/suncycle.html> [Přístup získán 15. 5. 2016].
- [25] *Informace o funkci DataHash* [Online] Dostupné z: <http://www.mathworks.com/matlabcentral/fileexchange/31272-datahash> [Přístup získán 15. 5. 2016].
- [26] *Popis webového rozhraní pro získání předpovědních dat ze služby PV Forecast* [Online] Dostupné z: http://www.pvforecast.cz/download/PV_Forecast_WEB_API_1.2.pdf [Přístup získán 15. 5. 2016].
- [27] *Department of Defense World Geodetic System 1984, Its Definition and Relationships with Local Geodetic Systems (Chapter 3.3)* [Online] Dostupné z: <http://earth-info.nga.mil/GandG/publications/tr8350.2/wgs84fin.pdf> [Přístup získán 15. 5. 2016].
- [28] *Přístup k datům z Univerzitního centra energeticky efektivních budov* [Online] Dostupné z: <http://147.32.99.42/vdv/vdv.php> [Přístup získán 15. 5. 2016].
- [29] Mathieu David, Hadja Maïmouna Diagne, Philippe Lauret: *Outputs and Error Indicators for Solar Forecasting Models* [Online] Dostupné z: https://ases.conference-services.net/resources/252/2859/pdf/SOLAR2012_0651_full%20paper.pdf [Přístup získán 15. 5. 2016].

- [30] J. Zhang, B.-M. Hodge, A. Florita, S. Lu, H. F. Hamann, V. Banunarayanan: *Metrics for Evaluating the Accuracy of Solar Power Forecasting* [Online] Dostupné z: <http://www.nrel.gov/docs/fy14osti/60142.pdf> [Přístup získán 15. 5. 2016].
- [31] ČSN CLC/TS 61836 *Solární fotovoltaické energetické systémy - Termíny, definice a značky* Brusel, březen 2012
- [32] *Persisitent variable* [Online] Dostupné z: <http://www.mathworks.com/help/matlab/ref/persistent.html> [Přístup získán 21. 5. 2016].
- [33] *PV Forecast* [Online] Dostupné z: <http://www.pvforecast.cz/> [Přístup získán 21. 5. 2016].
- [34] Jan Šulc, Martin Procházka, Marek Maška, Jan Včelák: *Návrh algoritmů pro získání předpovědi dopadajícího slunečního záření*, Alternativní zdroje energie 2016, Kroměříž, červen 2016
- [35] J. Včelák, P. Wolf, J. Šulc, M. Procházka, M. Maška, J. Kohout: *Využití předpovědi osvitů pro řízení hybridních fotovoltaických systémů*, Nekonvenční zdroje elektrické energie, Vyškov, květen 2016
- [36] Vyhláška č. 82/2011 Sb., Vyhláška o měření elektřiny a o způsobu stanovení náhrady škody při neoprávněném odběru, neoprávněné dodávce, neoprávněném přenosu nebo neoprávněné distribuci elektřiny