

**ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE**

**FAKULTA ELEKTROTECHNICKÁ**

Katedra měření



**Zpřesnění predikce slunečního osvětlení  
pomocí umělých neuronových sítí**

Diplomová práce

Vypracoval: Bc. Zbyněk Bambušek

Vedoucí práce: Ing. Petr Wolf PhD.

2017



## ZADÁNÍ DIPLOMOVÉ PRÁCE

Student: **Bc. Zbyněk Bambušek**

Studijní program: **Inteligentní budovy**

Název tématu česky: **Zpřesnění predikce slunečního osvětlení pomocí umělých neuronových sítí**

Název tématu anglicky: **Improving the Accuracy of Solar Irradiance Prediction Using Artificial Neural Networks**

### Pokyny pro vypracování:

Popište dostupné zdroje předpovědi počasí vhodné pro využití ve fotovoltaických systémech. Popište základní typy předpovědí dle časového horizontu a lokalizace a jejich typické zdroje dat a následné využití.

Definujte způsoby určení přesnosti předpovědi intenzity záření, určete přesnost vybraných zdrojů pro lokalitu centra UCEEB za historické období.

Využijte modul neuronových sítí v prostředí Matlab pro zpřesnění predikce vybraného nejvhodnějšího dostupného zdroje. Určete strukturu neuronové sítě, učící algoritmus a časové období pro učení tak, aby se dosáhlo co nejlepšího zpřesnění pro lokální predikci intenzity záření. Implementujte tuto službu na server.

Použijte též neuronovou síť pro predikování výkonu zvoleného fotovoltaického systému, porovnejte přesnost s konvenčním modelem přepočtu intenzity záření a teploty na výkon fotovoltaického systému.

### Seznam odborné literatury:

- [1] Staněk, K.: Fotovoltaika pro budovy. Praha: Grada, 2012.
- [2] Paulescu, M., Paulescu, E., et al.: Weather Modeling and Forecasting of PV Systems Operation, Springer-Verlag, London 2013
- [3] Šnorek, M.: Neuronové sítě a neuropočítače, ČVUT, Praha, 2002

Vedoucí diplomové práce: **Ing. Petr Wolf, Ph.D. (UCEEB ČVUT)**

Datum zadání diplomové práce: **1. února 2017**

Platnost zadání do<sup>1</sup>: **30. září 2018**

Prof. Ing. Jan Holub, Ph.D.  
vedoucí katedry



Prof. Ing. Pavel Ripka, CSc.  
děkan

V Praze dne 1. 2. 2017

<sup>1</sup> Platnost zadání je omezena na dobu tří následujících semestrů.

## Abstrakt

Hlavním cílem této diplomové práce je vytvořit algoritmus pro predikci osvitů fungující na principu neuronových sítí, jehož smyslem je nahradit stávající metodu založenou na lineární interpolaci, která je v současné době implementována ve službě PV Forecast.

Výstupem tohoto algoritmu je předpověď osvitů (úhrnná intenzita ozáření) na horizontální rovinu pro vybrané lokality v České republice ve  $\text{Wm}^{-2}$ . Součástí práce je řada metod pro zpracování a hledání parametrů sítě, vstupních a výstupních dat a dále srovnání s aktuální verzí služby a některými dalšími veřejně dostupnými službami.

V samotném závěru práce je demonstrován další potenciál neuronových sítí v pozměněné verzi algoritmu, která místo osvitů předpovídá přímo poskytovaný výkon za jednotku času na fotovoltaické elektrárně instalované v lokaci Unhošť.

## Klíčová slova

Forecast, Předpověď, Výkon, Osvit, Neuronová síť, Preprocessing, PV Forecast, MATLAB

## Abstract

The main aim of this thesis is to create an algorithm to predict irradiation that will function on the principle of neural networks. The purpose of this algorithm is to replace the method of prediction based on linear interpolation which is currently implemented in the service PV Forecast.

The output is a prediction of irradiation (total intensity of irradiation) to horizontal surface for selected localities in Czech Republic in  $\text{Wm}^{-2}$ . The work contains various methods for processing and finding parameters for the network, preliminary and output data, as well as a comparison with the current version of the service and other publicly available services.

In the last part of the thesis, further potential of neural networks is demonstrated in a tweaked version of the algorithm which instead of irradiation predicts the amount of energy generated in a set amount of time by the photovoltaic power plant in Unhošť.

## Keywords

Forecast, Prediction, Power, Irradiation, Neural Network, Preprocessing, PV Forecast, MATLAB

## 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 autora práce

## Poděkování

Zde bych rád poděkoval panu Ing. Petru Wolfovi Ph.D. za cenné rady, podporu a pomoc, které mi během mé práce poskytl.

Dále bych chtěl poděkovat všem členům týmu RP5 Univerzitního centra energeticky efektivních budov ČVUT, zejména Ing. Tomáši Baumeltovi, Ing. Marku Maškovi, Bc. Martinu Procházkovi.

V neposlední řadě bych rád poděkoval své rodině a přátelům.

# OBSAH

<b>SEZNAM OBRÁZKŮ .....</b>	<b>IIIX</b>	
<b>SEZNAM TABULEK .....</b>	<b>XI</b>	
1	MOTIVACE .....	1
2	ÚVOD.....	2
3	ÚVOD DO NEURONOVÝCH SÍTÍ .....	3
4	PRINCIP ALGORITMU PV FORECAST.....	6
5	POUŽÍVANÁ DATA .....	8
5.1	<i>Přepoččet časové značky na solární azimut a horizont.....</i>	<i>8</i>
5.2	<i>Měřená data .....</i>	<i>10</i>
5.3	<i>Dostupné zdroje predikce.....</i>	<i>11</i>
5.4	<i>Dostupné zdroje oblačnosti.....</i>	<i>12</i>
6	APARÁT UŽITÝ PRO VÝPOČTY CHYB.....	13
7	NAVRŽENÁ ŘEŠENÍ PRO PREPROCESSING DAT.....	15
7.1	<i>Implementace zpracování dat v prostředí MATLAB .....</i>	<i>15</i>
7.2	<i>Úpravy posunutí dat vůči realitě .....</i>	<i>15</i>
7.2.1	<i>Osvit.....</i>	<i>16</i>
7.2.2	<i>Oblačnost.....</i>	<i>20</i>
7.3	<i>Stanovení korelace dat dostupných zdrojů osvitů .....</i>	<i>24</i>
7.4	<i>Srovnání chyby sítě s různými zdroji osvitů .....</i>	<i>24</i>
8	AKTUÁLNÍ SÍŤ .....	26
8.1	<i>Trénování neuronové sítě.....</i>	<i>26</i>
8.2	<i>Hledání vhodné topologie sítě.....</i>	<i>31</i>
8.3	<i>Funkce pro vytvoření a trénování sítě .....</i>	<i>35</i>
8.3.1	<i>Dodatečný preprocessing dat před vstupem do sítě.....</i>	<i>35</i>
8.3.2	<i>Skripty pro vytvoření a trénink sítě.....</i>	<i>36</i>
8.3.3	<i>Funkce pro různá zpracování výsledků .....</i>	<i>36</i>
8.4	<i>Závěrečné srovnání zdrojů .....</i>	<i>36</i>
9	PREDIKCE VÝKONU NA FV ELEKTRÁRNĚ UNHOŠŤ .....	39
10	ZPRACOVANÉ VÝSTUPY PŘEDPOVĚDI OSVITU .....	40
11	ZPRACOVANÉ VÝSTUPY PŘEDPOVĚDI VÝKONU .....	43
12	ZÁVĚR .....	45
13	BUDOUCÍ PRÁCE .....	47
13.1	<i>Zpracování a stahování dat .....</i>	<i>47</i>
13.1.1	<i>Osvit.....</i>	<i>47</i>
13.1.2	<i>Oblačnost.....</i>	<i>47</i>
13.2	<i>neuronová síť – predikce osvitů.....</i>	<i>48</i>

13.3	<i>neuronová síť – predikce výkonu</i> .....	48
13.4	<i>Nowcasting</i> .....	48
	<b>VYSVĚTLIVKY</b> .....	<b>50</b>
	<b>SEZNAM POUŽITÝCH ZDROJŮ</b> .....	<b>51</b>
	<b>OBSAH PŘILOŽENÉHO MÉDIA</b> .....	<b>54</b>



## Seznam obrázků

Obrázek 1: Funkce používané uvnitř neuronu (aktivační funkce) .....	4
Obrázek 2: Schématické zobrazení neuronové sítě .....	4
Obrázek 3: Interpolační tabulka .....	6
Obrázek 4: Solární azimut a horizont, jak je definován v práci .....	8
Obrázek 5: Deklinace .....	9
Obrázek 6: Průběhy polohy slunce na obloze pro náhodný vzorek dat .....	10
Obrázek 7: Grafická interpretace korelací dvou vektorů dat .....	14
Obrázek 8: Výchozí posunutí dat vůči reálným hodnotám .....	16
Obrázek 9: Posunutí dat vůči reálným hodnotám pro měsíc únor .....	16
Obrázek 10: Posunutí dat vůči reálným hodnotám pro měsíc červenec .....	17
Obrázek 11: Data po stažení.....	17
Obrázek 12: Data po prvním kroku preprocesingu .....	17
Obrázek 14: Data po třetím kroku preprocesingu .....	18
Obrázek 13: Data po druhém kroku preprocesingu .....	18
Obrázek 15: Data na konci preprocesingu .....	18
Obrázek 16: Vychýlení dat vůči realitě na celých datech (původní) .....	19
Obrázek 17: Vychýlení dat vůči realitě na celých datech (po úpravách).....	19
Obrázek 18: Posun dat vůči realitě po úpravách, vypočtená korelace .....	20
Obrázek 19: Posun dat oblačnosti pro zdroj Energocentrum.....	21
Obrázek 20: Posun dat oblačnosti pro zdroj Forecast IO.....	22
Obrázek 21: Posun dat oblačnosti pro zdroj Noaa1 .....	22
Obrázek 22: Posun dat oblačnosti pro zdroj Noaa2 .....	23
Obrázek 23: Neuronová síť, stávající model.....	26
Obrázek 24: Srovnání (2.4.2016) .....	28
Obrázek 25: Srovnání (3.4.2016) .....	28
Obrázek 26: Příklad 1, výstupní graf z úlohy serveru pro interval 26.1.2017 – 2.10.2017 .....	32
Obrázek 27: Příklad 2, výstupní graf z úlohy serveru pro interval 2.10.2017-25.2.2017 .....	32
Obrázek 28: Příklad 3, výstupní graf z úlohy serveru pro interval 30.7.2016-29.8.2016.....	33
Obrázek 29: Příklad 4, výstupní graf z úlohy serveru pro interval 27.11.2016-27.12.2016....	33
Obrázek 30: Hledání vhodného počtu neuronů pro předpověď osvitů ČVUT UCEEB .....	34
Obrázek 31: Schéma algoritmu na principu neuronových sítí .....	35
Obrázek 32: Hledání vhodného počtu neuronů pro předpověď výkonu Unhošť.....	39

Obrázek 33: Srovnání predikcí, $MRSE_{rel}$ pro denní sumy .....	40
Obrázek 34: Srovnání predikcí, $MRSE_{rel}$ pro hodinové hodnoty.....	41
Obrázek 35: Srovnání predikcí, $MAE_{rel}$ pro hodinové hodnoty.....	41
Obrázek 36: Srovnání predikcí, $MBE_{rel}$ pro hodinové hodnoty.....	42
Obrázek 37: Vyhodnocení predikce výkonu $RMSE_{rel}$ .....	43
Obrázek 38: Vyhodnocení predikce výkonu $MAE_{rel}$ .....	43
Obrázek 39: Vyhodnocení predikce výkonu $MBE_{rel}$ .....	44
Obrázek 40: Náhled výstupu sítě na serveru .....	45

## Seznam tabulek

Tabulka 1: Korelace predikcí osvitů po preprocesingu.....	24
Tabulka 2: Srovnání zdrojů predikcí při stejných podmínkách .....	24
Tabulka 3: Porovnání výběrů vhodné sítě (RMSE suma z jednotlivých měsíců).....	30
Tabulka 4: Porovnání výběrů vhodné sítě (RMSE <sub>rel</sub> suma z jednotlivých měsíců) .....	31
Tabulka 5: Srovnání zdrojů .....	37
Tabulka 6: Srovnání zdrojů pro Noaa2 .....	38

# 1 Motivace

Smyslem této diplomové práce je optimalizace predikce osvitů, které bude využíváno pro predikci výkonu solární elektrárny a případné obchodování s energií. Jde o pokračování projektu I [1] a na něj navazujícího projektu II [2] předchozích prací vedených stejně jako tato pod výzkumným ústavem ČVUT UCEEB<sup>1</sup>.

Přestože zastoupení solární energie v ČR mezi lety od roku 2012 spíše stagnuje [3], je procentuální zastoupení fotovoltaiky vysoké. Instalovaný výkon se totiž po předchozím boomu ustálil na 9,5 % [4]. Základní nevýhodou proměnných zdrojů energie, kam se řadí i fotovoltaika, je značná výkonová nestabilita a nákladná nebo nedostatečná technologie ukládání vyrobené energie. Zejména tyto důvody kladou požadavky na předpovědi výkonu elektráren. Přebytky a nedostatky v síti se musí trvale vyrovnávat, jinak dochází k ztrátám na zisku, může dojít k výpadku nebo dokonce poškození přenosové sítě.

Rozdíl špičkových výkonů mezi dvěma extrémními dny může v celostátním měřítku dosahovat hodnoty jednotek GW. Obecně lze říci, že využít přebytky je podstatně snazší, tím pádem i levnější, pokud se o nich ví předem.

Další praktický význam mají predikce pro obchodníky s elektřinou, kde je potřeba neustále udržovat nabídku a poptávku energie v rovnováze.

---

<sup>1</sup> Univerzitní centrum energeticky efektivních budov

## 2 Úvod

Cílem práce je vytvoření algoritmu pro předpověď osvitů na horizontální rovinu, který je v současné době poskytován jako zpoplatněná služba [29] pod názvem PV Forecast (photovoltaic forecast) ústavem ČVUT UCEEB.

Nyní, tedy před navržením algoritmu na principu neuronových sítí, služba již funguje s jednoduchým algoritmem založeným na lineární interpolaci, který bude popsán dále.

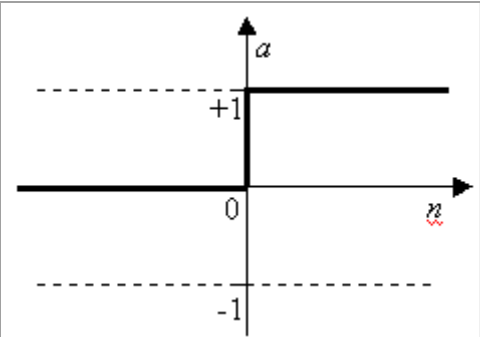
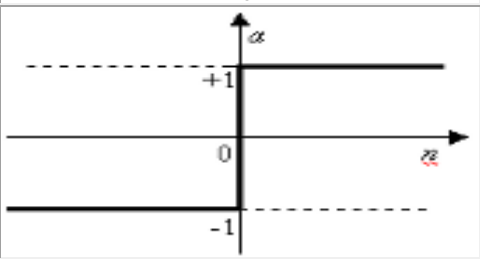
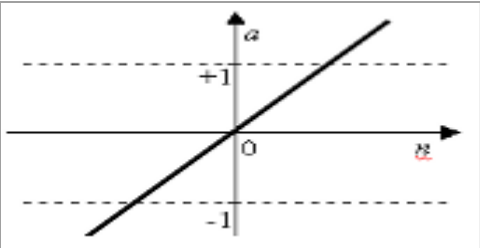
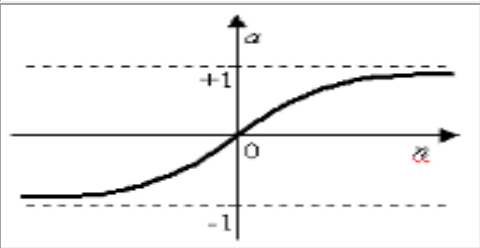
Dva již zmíněné předcházející projekty jsou k práci přiloženy. První projekt je spíše seznamovací, projekt druhý navrhuje funkční kostru algoritmu a demonstruje, že dosahuje lepších výsledků. Součástí této navazující práce je zejména odladění již existujícího algoritmu a jeho rozšíření o nové skripty a zpracování dat.

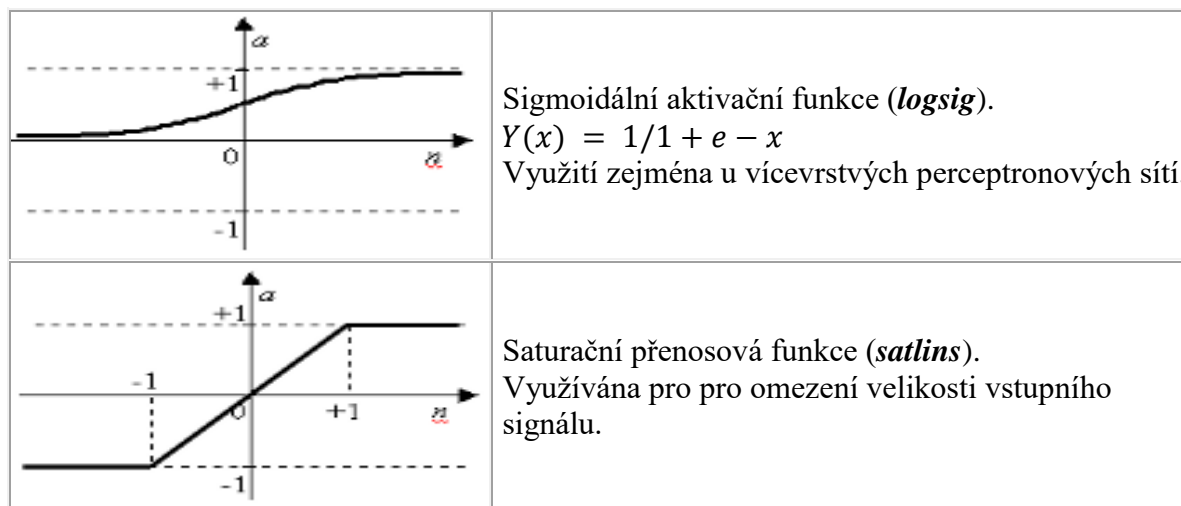
Tuto práci lze obecně rozdělit do několika následujících bodů:

- 1) Úvod do problematiky, stručný popis principů obou algoritmů.
- 2) Představení použitých zdrojů dat, způsob jejich úpravy a zavedení nového preprocesingu dat.
- 3) Postupy pro odhalení a vyhodnocování chyb.
- 4) Představení konkrétní neuronové sítě použité v novém algoritmu, její implementace v prostředí MATLAB, metody odladění některých jejích vlastností.
- 5) Srovnání algoritmu na bázi neuronové sítě a původního algoritmu, srovnání s konkurenčními zdroji.
- 6) Nastínění využití neuronových sítí pro předpověď výkonu na konkrétní aplikaci v elektrárně Unhošť. Srovnání jednostupňové metody a metody, kdy je nejprve předpovězen osvit a ten je následně přepočítán na výkon.
- 7) Návrh možných zlepšení a dalšího směřování práce.

### 3 Úvod do neuronových sítí

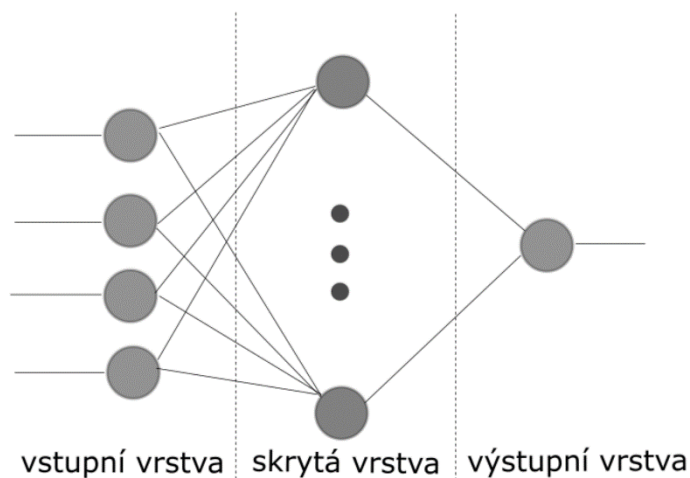
Neuronová síť je výkonná početní metoda, která vytváří jeden nebo více výstupů na základě pevně daného počtu vstupních a výstupních parametrů. Základní princip je odpozorovaný z biologických neuronů. Tak jako v biologii i v neuronové síti je základním prvkem neuron (percepton). Při průchodu signálu neuronem je signál změněn. To, jak se signál mění, definuje aktivační funkce, tedy funkce, která je v neuronu obsažena. Obvykle se v případě neuronových sítí používají funkce z obr. 1. [5] V této práci je používá defaultním nastavení, tedy hyperbolický tangens v první vrstvě (na obr. 2 je to skrytá vrstva), lineární aktivační funkce ve druhé (na obr. 2 je to výstupní vrstva).

	<p>Funkce jednotkového skoku (<i>hardlim</i>). Standardní výstup nenabuzeného neuronu je hodnota <math>0</math>, po nabuzení <math>-1</math>.</p>
	<p>Aktivační funkce signum (<i>hardlims</i>). <math>Y(x) = \text{sign}(x) = 1</math> pro <math>x &gt; 0</math> a <math>Y(x) = \text{sign}(x) = -1</math> pro <math>x &lt; 0</math>.</p>
	<p>Lineární aktivační funkce (<i>purelin</i>). <math>Y(x) = x</math></p>
	<p>Aktivační funkce hyperbolický tangens (<i>tansig</i>). Využití u Elmanových neuronových sítí.</p>



Obrázek 1: Funkce používané uvnitř neuronu (aktivační funkce)

Spojením několika neuronů vzniká neuronová síť. Síť zvolená pro naši aplikaci má tři vrstvy. Vrstvu vstupní, skrytou vrstvu a vrstvu výstupní. Jde o model, který se používá nejčastěji, je jednoduchý a je vhodný pro aproximaci nelineární funkce. Na obr. 2 je zobrazena neuronová síť, která má 4 vstupní neurony a jeden výstupní. Každý z neuronů neuronové sítě přijímá vstupy, které může měnit na očekávané výstupy.



Obrázek 2: Schématické zobrazení neuronové sítě

Pouhá neuronová síť pro předpověď nestačí, je potřeba ji nejdříve natrénovat. Při trénování dochází k nastavování parametrů aktivační funkce každého neuronu tak, aby reakce na požadované vstupy odpovídala požadovaným výstupům.

Pro trénink sítě byla užitá data naměřená pro lokalitu ČVUT UCEED. Původní myšlenka byla, že společná část dat bude použita jako referenční pro komparaci původního algoritmu služby

PV Forecast<sup>ii</sup> a jeho nové verze. Nakonec se ale ukázalo, že po nových úpravách existují lepší zdroje vstupních dat a aby bylo srovnání autentické, je při porovnávání PV Forecast založen na stejných datech, jaké doposud používá. Kromě toho má neuronová síť na vstupu více typů dat, takže srovnání na úplně identických datech by ani tak nebylo realizovatelné.

---

<sup>ii</sup> Dále bude spojení „PV Forecast“ vždy označovat původní algoritmus, který má být sítí nahrazen.



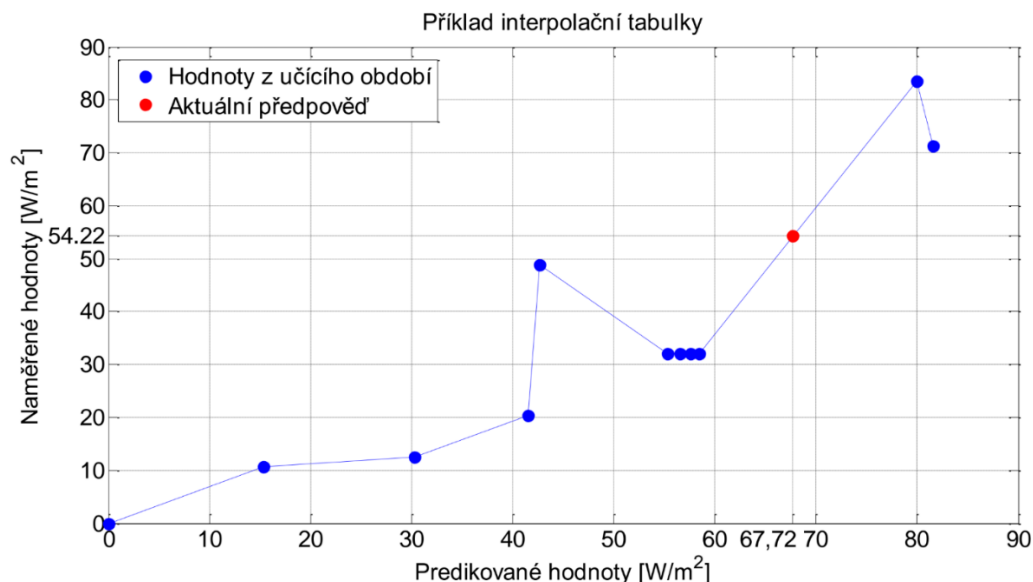
## 4 Princip algoritmu PV Forecast

Než bude blíže pokračovat v řešení problému pomocí neuronových sítí, je vhodné ještě krátce odbočit a zmínit se o stávajícím algoritmu, který má být sítí nahrazen.

Původní algoritmus je v současné době použit k predikci na 48 hodin dopředu a jako vstupy používá jen cizí zdroj predikce osvitů a data změřená na daném místě. Pro srovnání s algoritmem založeným na neuronové síti byla vybrána lokalita Buštěhrad (ČVUT UCEEB). Jednotlivé zdroje pro předpovědi jsou stahovány ze zdrojů Národní úřad pro oceán a atmosféru, Pacific Islands Ocean Observing System, Medard, Energocentrum PLUS, s.r.o a Solar irradiation database, naměřená data jsou ze zdrojů ČVUT UCEEB, Domanín, Nový Lískovec, Jeseník (budova Fénix Group).

Podrobnější analýza dat zdrojových predikcí uvažovaná pro tento algoritmus je popsána v bakalářské práci *Predikce krátkovlnného slunečního záření pro malé fotovoltaické systémy*. [6] Každý ze zdrojů byl statisticky zpracován a zdroje byly seřazeny od nejlepšího po nejhorší. Pro predikci se potom užije první dostupný nejlepší zdroj.

Data jsou předzpracována a v případě potřeby (např. pro zdroj Hawaii se získají data v trojhodinových intervalech) jsou data interpolována a dále už se pracuje s jednotným formátem dat, která jsou vzorkována po hodině.



Obrázek 3: Interpolační tabulka, převzato z práce Matrina Procházky [6]

Principem je proces učení, kdy jsou historická data z posledních dvou týdnů strukturována do interpolačních tabulek (viz obr. 3). Každé hodině tak odpovídá jedna tabulka složená

z maximálně čtrnácti historických vzorků. V případě, že jsou hodnoty nevhodně rozmístěny, se problematické hodnoty nahradí jednou společnou.

Z těchto tabulek se potom přímo získává požadovaná hodnota. Naměřená „budoucí“ hodnota se jednoduše stanoví průmětem známé predikce jednoho ze zdrojů přes graf na osu y. Tento postup se uplatní pro každou hodinu (tabulku) a výsledkem je predikce na 48 hodin dopředu. Výstupní data jsou potom filtrována a odeslána na server.

## 5 Používaná data

V této kapitole budou stručně představeny jednotlivé dostupné zdroje dat použitých v práci. Rozbor některých z nich je podrobněji zpracován v práci pana Procházky<sup>iii</sup> [6], ze které také těží rozbor zdrojů dat v podkapitolách 5.3 *dostupné zdroje predikce* a 5.4 *dostupné zdroje oblačnosti* této práce.

Typy vstupních dat použitých pro vstup neuronové sítě:

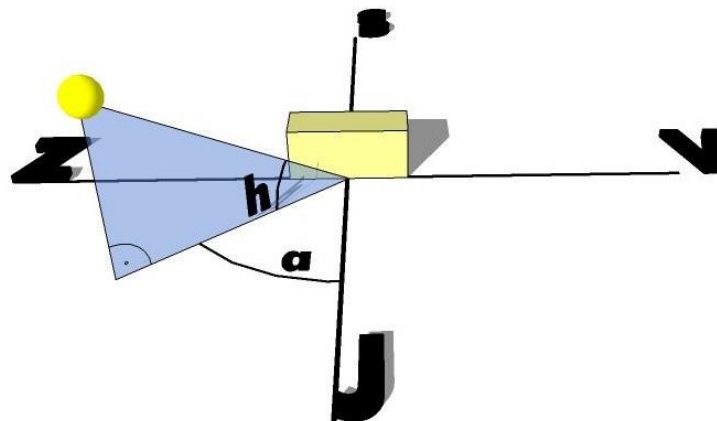
- časová značka, která je přepočtena na pozici Slunce.
- predikce z cizího zdroje.
- predikce oblačnosti

Výstupní data užitá pro trénování sítě:

- reálné naměřené hodnoty intenzity osvitů v ČVUT UCEEB
- reálné naměřené hodnoty výkonu v Unhošti

### 5.1 Přepočet časové značky na solární azimut a horizont

Vzhledem k faktu, že datum a čas jsou nespojité veličiny, je dobré je pro síť předzpracovat. Nespojitost se běžně projevuje v hodinách, dnech i měsících a pro síť by bylo obtížné tato data při trénování „pochopit“. To je důvodem, proč místo časové značky vstupuje do trénovacího algoritmu vstupuje pozice slunce řídicí se popisem na obr. 4.

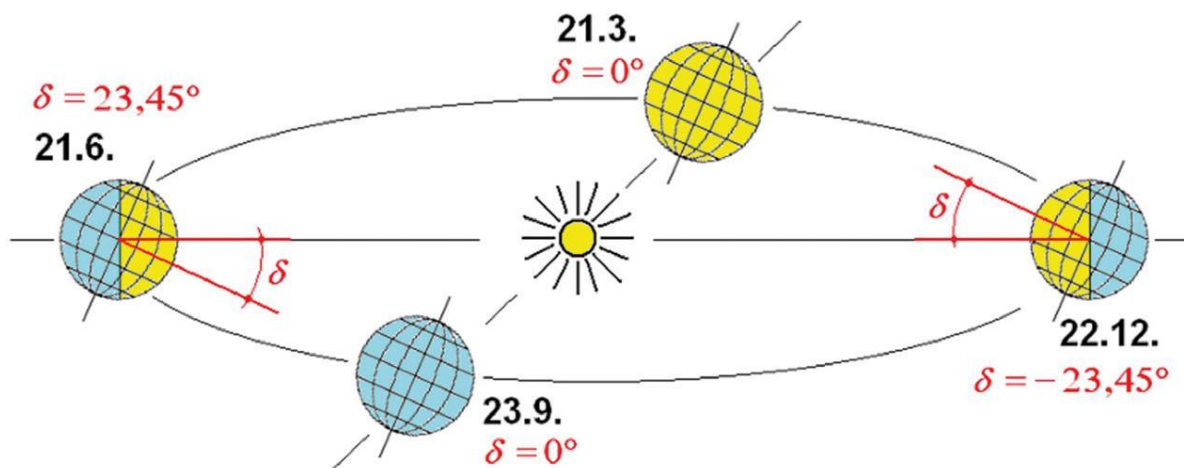


Obrázek 4: Solární azimut a horizont, jak je definován v práci

<sup>iii</sup> Viz kap 2.

Pro výpočet polohy slunce je potřeba nejdříve určit **deklinaci** [7] (obr. 5) podle vztahu:

$$\delta = 23,45^\circ \sin\left(360 \frac{284+n}{365}\right), \text{ kde } \delta \text{ značí deklinaci a } n \text{ pořadí dne v roce.}$$



Obrázek 5: Deklinace

Následuje **výška slunce nad obzorem** (tj. solární horizont) [7] (obr. 5):

$$\tau = 15^\circ(ST - 12)$$

$$h = \sin^{-1}(\sin \delta \sin \phi + \cos \delta \cos \phi \cos \tau),$$

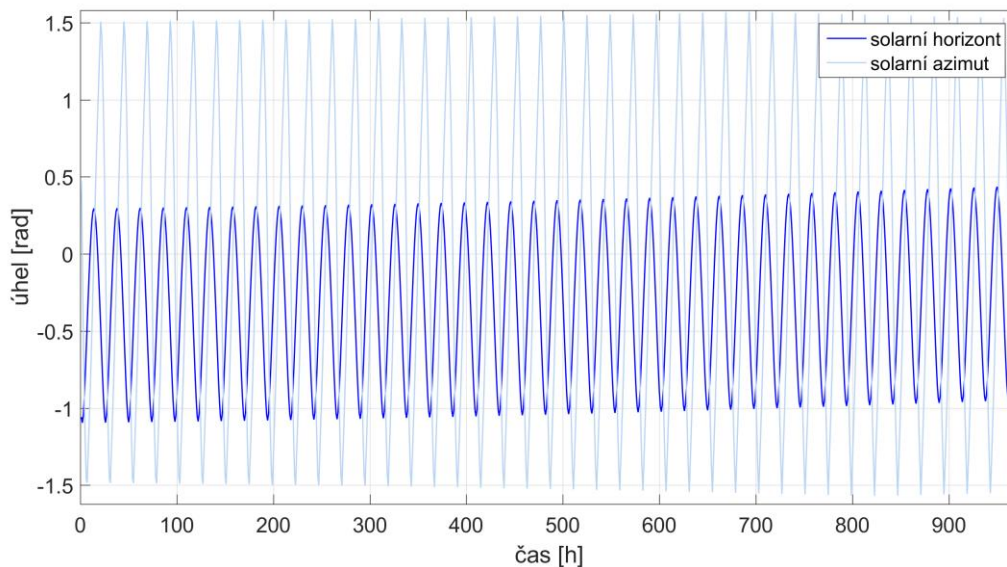
kde  $h$  je výška slunce nad obzorem,  $\phi$  je 50.1566907 zeměpisná šířka budovy ČVUT UCEEB.  $\tau$  je sluneční časový úhel a  $ST$  je sluneční čas. Jde o čas takový, aby ve 12 hodin (zimního času) bylo Slunce přesně v nadhlavníku, což se v algoritmu řeší posunem o  $-\frac{\varphi}{15}$ , kde  $\varphi$  značí zeměpisnou délku pro ČVUT UCEEB tedy 14.1695708.

A nakonec **solární azimut** [7] (obr. 5):

$$a = \sin^{-1}\left(\frac{\cos \delta}{\cos h} \sin \tau\right),$$

kde  $a$  je sluneční azimut.

Pomocí výšky slunce nad obzorem je možné vyřadit noční hodnoty. Při úpravě dat jde spíše o orientační vyřazení nočních hodnot, která usnadňuje průběh algoritmu. (Při učení se vyřazují nulové noční hodiny.)



Obrázek 6: Průběhy polohy slunce na obloze pro náhodný vzorek dat

## 5.2 Měřená data

Cílem této práce je vytvoření algoritmu, který bude možné v budoucnu rozšířit na celou Českou republiku. Pro vyhodnocení algoritmu jsou ovšem dostačující už měřená data z lokality Buštěhrad (ČVUT UCEEB). Data z místního zdroje jsou měřena v minutových intervalech. Konečný algoritmus pracuje v intervalech hodinových, proto je nutné data nejdříve upravit. K synchronizaci vzorkování je použit aritmetický průměr, kde uvažovaná hodina leží uprostřed intervalu. Průměruje se tedy 30 minut před a 30 minut po dané hodině.

Do budoucna jsou k dispozici ještě další zdroje, a při implementaci na server se také uplatní. Půjde o stejné zdroje jako pro algoritmus PV Forecast, více je možné se o nich dozvědět v práci pana Martina Procházky [6].

V poslední části práce bude rozebrán možný přínos předpovědi výkonu. K tomuto účelu jsou použita data o poskytnutém výkonu ze zdroje Unhošť, který je necelých 20 km od ČVUT UCEEB. (Pro účel práce je možné použít některá data pro ČVUT UCEEB.) Data o výkonu nejsou určena veřejnosti, nebudou tedy na přiloženém médiu. Data z Unhoště jsou od určitého data stahována v pětiminutových intervalech, do té doby po patnáctiminutových. Opět se pro přepočítání používá aritmetický průměr, patnáctiminutová data jsou nejdříve interpolována na pětiminutová.

### 5.3 Dostupné zdroje predikce

Z hlediska práce s daty jsou mnohem zajímavější predikce z jiných zdrojů, ze kterých se u algoritmu vychází. Pro každou lokalitu je dostupných více zdrojů a zejména do získání určitého všeobecného přehledu se s daty velmi obtížně pracuje. Data jsou pro různé lokace, jsou stažena v různých formátech i s různým vzorkováním. Na rozdíl od měřených dat jsou také zatížena nezanedbatelnými chybami.

- Pacific Islands Ocean Observing System (dále Hawaii) [20, 21]

Zdroj, označovaný také jako Hawaii, byl v práci pana Procházky, který je autorem původního algoritmu vyhodnocen jako nejvhodnější. Proto je použit v původním algoritmu jako výchozí.

Zdroj Hawaii poskytuje data pro celý svět v intervalu  $0.5^\circ$  zeměpisné šířky a délky. Data jsou stahována ve formátu *csv* se vzorkováním po 3 hodinách na 120 hodin dopředu. A jsou aktualizována každou hodinu. Před použitím je potřeba je značně upravit, protože mají kladné hodnoty i v nočních hodinách. K úpravě tohoto typu je použita polynomiální a lineární interpolace a zúžení dat vzhledem k časové ose podle aktuálního východu a západu. Data jsou veřejně dostupná.

- Medard [20]

Zdroj Medard poskytuje předpovědi ve 2 pásmech. Pro Českou republiku jde o hustotu pokrytí ve čtvercích území o straně 3 km pro vzorkování 2 dny dopředu, dále je dostupná předpověď pro 3 dny dopředu ale už jen při pokrytí čtverci o straně 9 km. Predikce je v hodinových intervalech. Předpověď je aktualizována vždy v 0,6,12 a 18 hodin. Zdroj slouží pro testovací účely.

- Národní úřad pro oceán a atmosféru [11,12,13]

- GEFS Reforecast v2 Server (dále Noaa1) [14]

Poskytuje data pro celý svět v intervalu  $1^\circ$  zeměpisné šířky a délky. Data jsou ve formátu *grib2*<sup>10</sup> se vzorkováním po 3 hodinách na 3 dny dopředu. Aktualizace probíhá jednou za den. Zdroj je dostupný jen pokud je uveden.

- ERDDAP NOAA Server (dále Noaa2) [15]

Poskytuje data pro celý svět v intervalu  $0,25^\circ$ ,  $0,5^\circ$  nebo  $1^\circ$  zeměpisné šířky a délky. Data jsou k dispozici ve formátu *grib2* [10] se vzorkováním po 3 hodinách na 2 dny dopředu. Predikce je přepočítána každých 6 hodin. Jde o veřejně dostupný zdroj.

- Solar Radiation Database (dále Soda) [16, 17]

Předpověď je pro celý svět v intervalu 1° zeměpisné šířky a délky. Data jsou ve formátu *csv*, se vzorkováním po 1 hodině. Aktualizována jsou jednou denně. Zdroj smí být použit jen pro osobní účely. V současné době data z tohoto zdroje už nejsou stahována. Došlo k zásadním změnám poskytovaných služeb.<sup>iv</sup>

## 5.4 Dostupné zdroje oblačnosti

Jedním ze zdrojů pro neuronovou síť je zakrytí oblohy. K dispozici jsou v současné době 4 zdroje, z nichž jeden bohužel nemůže být řádně vyhodnocen, vzhledem ke krátké historii dat. Nedostatky těchto zdrojů jsou podobné jako v případě dostupných zdrojů predikce.

- Noaa1 a Noaa2 [11,12,13,14,15]

Stejně jako predikci osvitů poskytují tyto služby i predikci řady dalších veličin včetně oblačnosti. Data jsou ve stejném tvaru jako v případě osvitů.

- Darksky (dále Forecast IO) [8]

Darksky je služba specializovaná na předpovídání počasí s přesností pod jednu minutu (To platí pro predikci na 1 hodinu dopředu). Služba poskytuje data ve formátu *json* [9] za účelem snadného zpracování v aplikacích. Tak jako u ostatních služeb předpovídá celou řadu dalších veličin. V provozu je od roku 2011. Je schopná poskytovat i data na 60 let zpětně.

Jde obecně o data o počasí pro celý svět. Predikce oblačnosti je na 2 dny dopředu, data jsou vzorkována po jedné hodině a jsou aktualizována každou hodinu. Služba je placená.

- ENERGOCENTRUM PLUS, s.r.o. (dále Energo centrum) [18,19]

Pro testovací účely je k dispozici předpověď pro Prahu na 72 hodin dopředu se vzorkováním po jedné hodině. Data jsou poskytována v 11:00 a 23:00 ve formátu *csv*. ENERGOCENTRUM PLUS poskytuje data jen pro testovací potřeby.

---

<sup>iv</sup> Jde tedy o zastaralé informace, které ale korespondují s v práci ještě užitými daty.

## 6 Aparát užitý pro výpočty chyb

Jako směrodatné pro vyhodnocení dat a výsledků zlepšení predikce sítí se uvažuje rozdíl v chybách mezi srovnávanými vektory. Jako referenční hodnoty jsou uvažovány hodnoty naměřené na stanici ČVUT UCEEB, protože jejich chybovost je ve srovnání s chybovostí zdrojů dat zanedbatelná.

➤ V rámci této práce se pracuje s následujícími chybami [23]:

1.  $RMSE = \sqrt{\frac{1}{N} \sum_i^N (x_{pred,i} - x_{target,i})^2}$  tj. směrodatná odchylka,
2.  $MBE = \frac{1}{N} \sum_i^N (x_{pred,i} - x_{target,i})$  tj. střední chyba uvažující záporné hodnoty,
3.  $MAE = \frac{1}{N} \sum_i^N |x_{pred,i} - x_{target,i}|$  tj. střední absolutní chyba.

➤ Dále z nich odvozené relativní chyby [23]:

1.  $RMSE_{rel} = \frac{\sqrt{\frac{1}{N} \sum_i^N (x_{pred,i} - x_{target,i})^2}}{\bar{x}_{target}}$
2.  $MBE_{rel} = \frac{1}{N} \sum_i^N \frac{(x_{pred,i} - x_{target,i})}{\bar{x}_{target}}$
3.  $MAE_{rel} = \frac{1}{N} \sum_i^N \frac{|x_{pred,i} - x_{target,i}|}{\bar{x}_{target}}$

Ve výše uvedených vztazích  $x_{pred,i}$  značí predikovanou hodnotu pro vzorek  $i$ ,  $x_{target,i}$  značí reálnou naměřenou hodnotu pro vzorek  $i$  a  $\bar{x}_{target}$  je aritmetický průměr ze všech reálných vzorků. Obecně je v celé práci nejvíce využívána směrodatná odchylka.

➤ Korelace:

Korelace je metoda, která stanovuje závislost mezi dvěma procesy, či veličinami. Pokud je výsledná korelace nenulová, znamená to, že mezi zkoumanými objekty může být vztah. Korelace neříká nic o kauzalitě, nabývá hodnot z intervalu  $< -1; 1 >$ , takže lze posoudit přímou a nepřímou úměrnost.

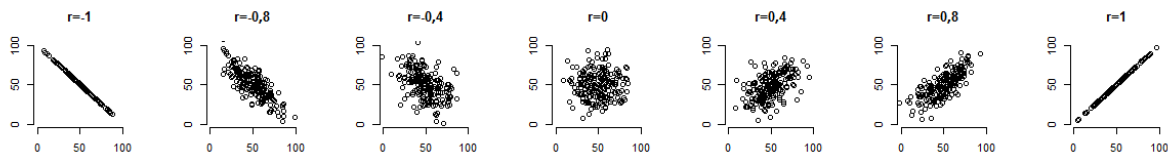
Obvykle se uvažuje jedna z následujících metod korelace [24].

- Pearsonův korelační koeficient: Obvykle slouží k měření statické závislosti lineárních dat, přesto ukazuje nějakou informaci o datech. V některých případech je použit.
- Kendallův korelační koeficient: Projevuje se větší citlivostí na některé nelineární vztahy. Rozdíl mezi pravděpodobností, že hodnoty dvou proměnných jsou ve stejném



pořadí, a pravděpodobností, že ve stejném pořadí nejsou. Pro naši aplikaci se citlivost neprojevila.

- Spearmanův korelační koeficient: Zachycuje monotónní vztahy mezi proměnnými nikoliv pouze lineární. Vzhledem k nelinearitě dat se ukázal jako nejvhodnější.



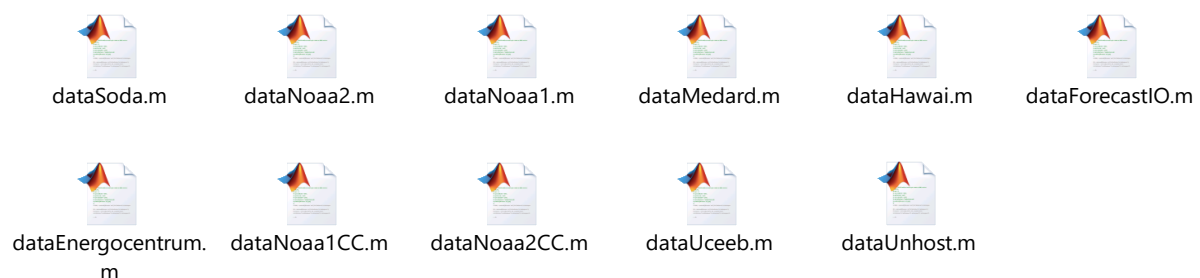
Obrázek 7: Grafická interpretace korelaci dvou vektorů dat, převzato z <http://www.wikiskripta.eu>

## 7 Navržená řešení pro preprocessing dat

Po seznámení se základním aparátem pro výpočet chyb dat již nic nebrání v rozebrání základních aplikací, které je potřeba provést, abychom data dostali do pro algoritmus použitelné podoby.

### 7.1 Implementace zpracování dat v prostředí MATLAB

Pro předpřipravení dat a jejich stahování ze zdrojů se užívá řada funkcí, které data vrací ve formátu vhodnějším pro další zpracování neuronovými sítěmi než původní funkce ze služby PV Forecast. Jde o následující funkce, jejichž názvy jsou poměrně přímočaré:



Některé z funkcí obsahují ještě volání podfunkcí, které obvykle obsahují převzatou část kódu ze stahování dat navržené původně pro PV Forecast. Tyto části kódu byly povětšinou rozsáhle upraveny a obvykle jde o pomocné funkce, které např. překládají speciální formáty dat. Proto zde tyto skripty nebudou uvedeny, nicméně budou přiloženy na médiu přiloženém k této práci.

### 7.2 Úpravy posunutí dat vůči realitě

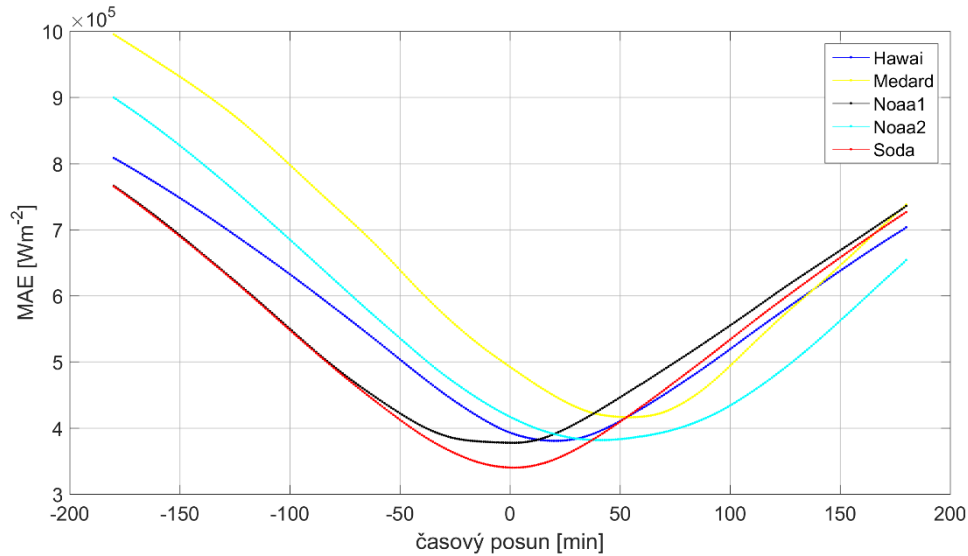
Vzhledem ke stavu stahovaných dat se otvírá řada možností, jak při zpracovávání dat může dojít k jejich posunutí. Není nic neobvyklého, že data jsou podávána s časovými značkami neodpovídajícími realitě, může dojít k malému posunu díky zaokrouhlení zeměpisné šířky/výšky, změně času apod. Zásadnější posun je možné detekovat i ve velmi nepřesných datech, jaká jsou k dispozici.

Pro demonstrativní odhalení posunu je nejprve užita interpolační metoda se zkratkou *pchip*<sup>v</sup> pro interpolaci dat na minutové intervaly. Poté se data posouvají vůči naměřeným datům pro úzké okolí, pro účely demonstrace 180 minut na obě strany a počítá se chyba dat, například MAE (viz. 6 Aparát užitý pro výpočty chyb).

<sup>v</sup> Jde o nelineární metodu pro interpolaci dat, více v dokumentaci k funkci v MATLABU *interp1.m*<sup>25</sup>.

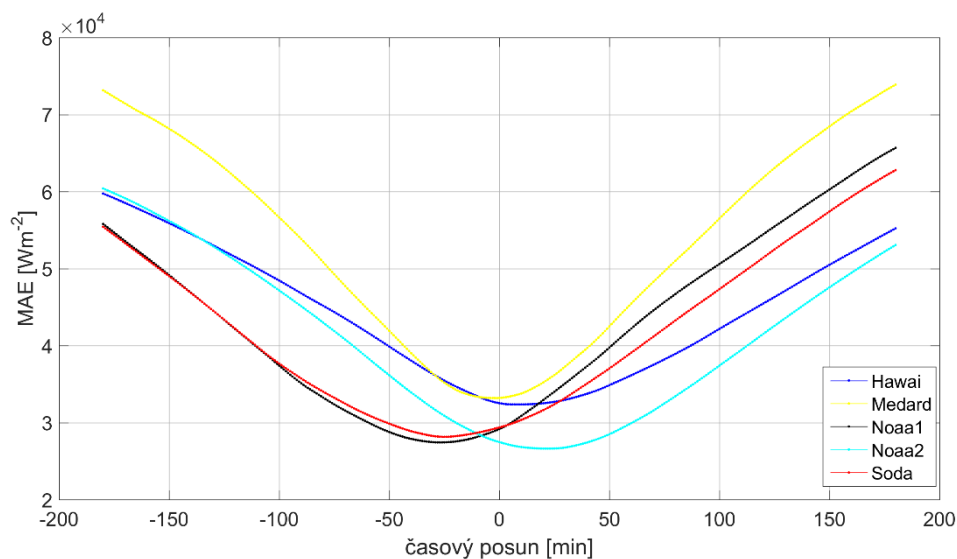
### 7.2.1 Osvit

Pro zdrojové predikce osvit je původní výsledek této metody provedený na datech stažených od 1.1.2016 do 21.3.2017 zobrazen na obr. 8.

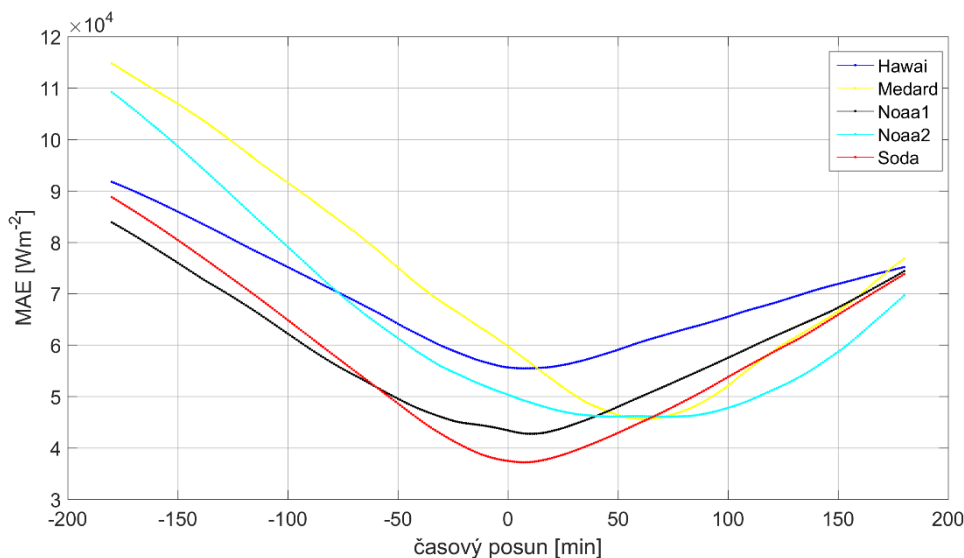


Obrázek 8: Výchozí posunutí dat vůči reálným hodnotám

Pochopitelně pro celý rok graf jen ukazuje, že zvažování posunu má smysl. Obecně budou mít letní měsíce na výsledek větší vliv než měsíce zimní, protože sluneční osvit má v těchto měsících mnohem vyšší hodnoty. Přitom letní měsíce mohou vlivem změny času na letní mít s touto změnou spojené posunutí. Grafy z obr. 9 a 10 ukazují situaci pro měsíce únor a červen. Při srovnání je patrná odlišnost vůči posunu celých dat. V červnu se pro některé ze zdrojů projevuje posun času.

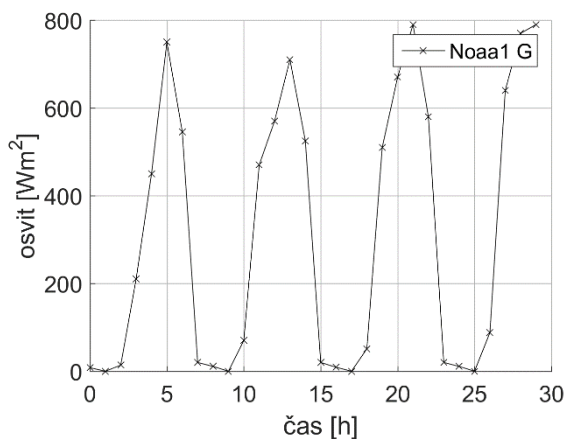


Obrázek 9: Posunutí dat vůči reálným hodnotám pro měsíc únor

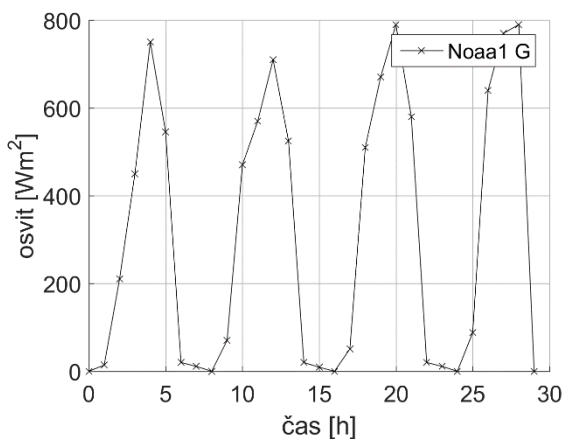


Obrázek 10: Posunutí dat vůči reálným hodnotám pro měsíc červenec

Na základě závislostí zjištěných z uvedených grafů a nedostatků původního kódu z projektů I a II., byl přepracován preprocessing dat ze solárního osvitů. Obecně šlo o následující změny, ačkoliv vždy nebylo potřeba uplatnit všechny. Na grafech doplňující následující text o preprocessingu dat je vývoj osvitů pro 15.6.2016 až 18.6.2016 pro zdroj Noaa1, který obsahuje všechny nově přidání části zpracování.



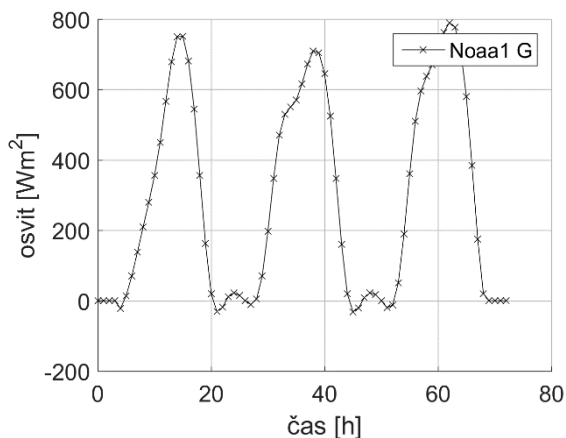
Obrázek 11: Data po stažení



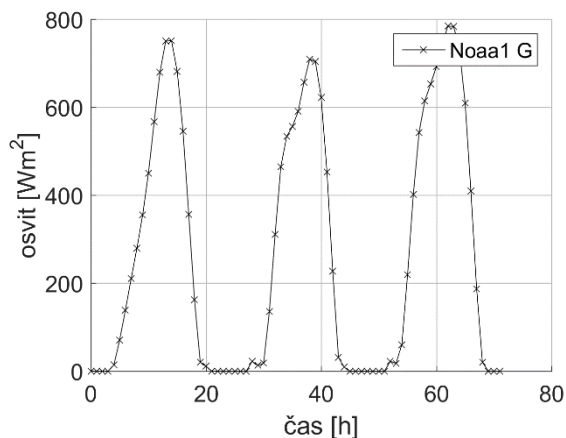
Obrázek 12: Data po prvním kroku preprocessingu

Při stažení dat, jsou časové značky pravděpodobně nastaveny vzhledem k času místa zdroje. To v praxi znamená, že zejména pro americké zdroje v létě, kdy jsou dny dlouhé, zapadá slunce kolem 2:00. Na informaci o konkrétní hodině sice není brán ohled, ale pro další postup je potřeba posunout data tak, aby celá byla v příslušném dni. Typicky stačil posun o 3 hodiny, což odpovídá jednomu vzorku viz grafy z obr 11 a 12.

Jak bylo několikrát zmíněno, data jsou často vzorkována po třech hodinách, místo hodinových intervalů, které mají být výstupem služby PV Forecast. Bylo tedy potřeba vyinterpolovat neznámé hodnoty pomocí nelineární metody. Na obr. 13 je vidět i chyba způsobená interpolací, zde konkrétně obloučky mezi dny.

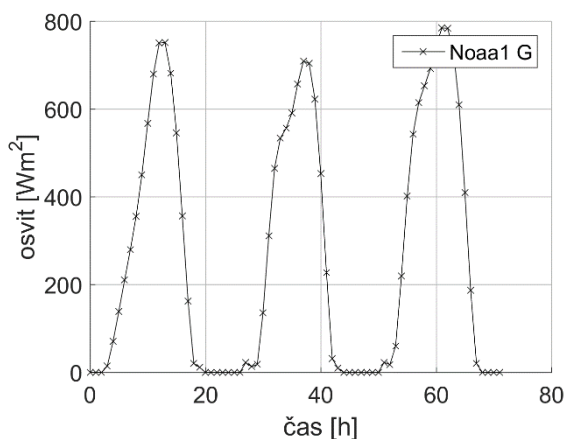


Obrázek 14: Data po druhém kroku preprocessingu



Obrázek 13: Data po třetím kroku preprocessingu, je patrné, že chyba interpolace nebyla dokonale opravena

Za první interpolací následuje další interpolace, která tentokrát zužuje tvar predikce viz rozdíl mezi grafy na obr. 13 a 14. Zároveň se v tomto kroku užívá převzaté funkce pro určení východu a západu slunce *suncycle.m*. A nenulová data jsou tedy zúžena na časový interval mezi východem a západem a ostatní hodnoty jsou nahrazeny nulami.



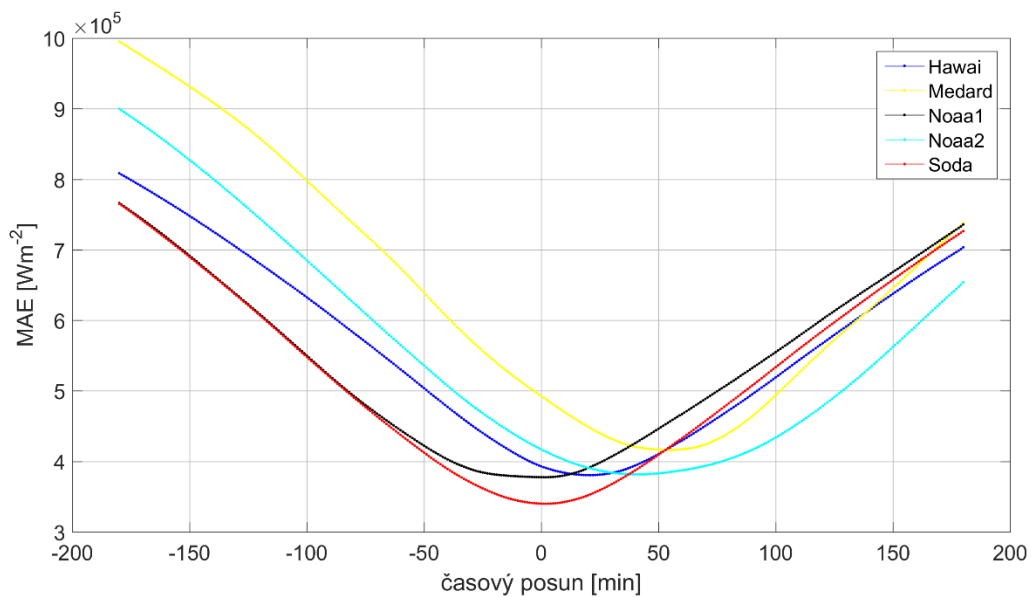
Obrázek 15: Data na konci preprocessingu

Poslední krok je posun času ze zimního na letní viz obr. 15.

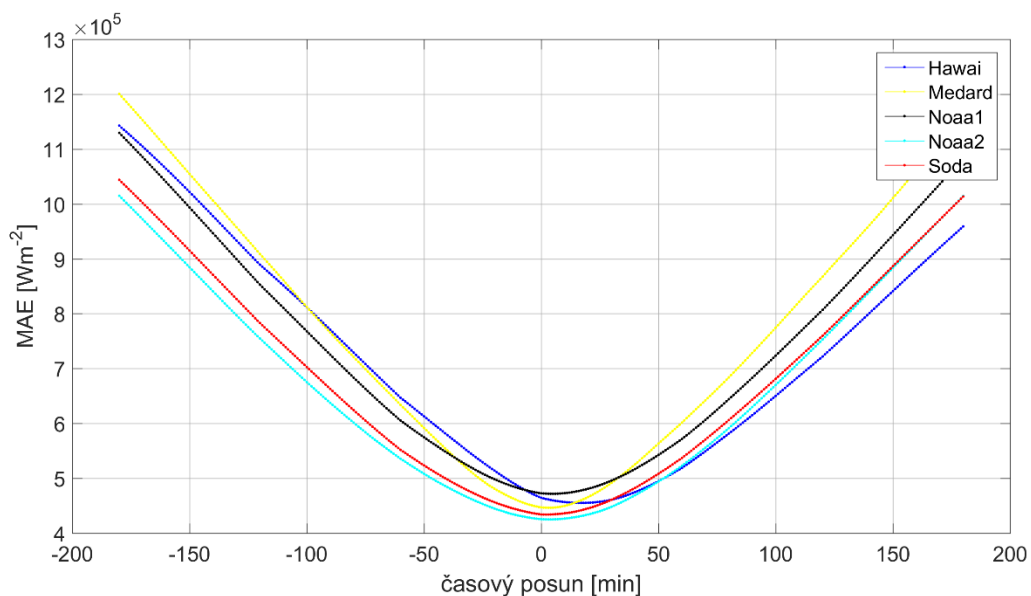
Při těchto postupech byl důraz na respektování očekávaného chování, tedy nebyly zahrnuty žádné „magické“ konstanty, ačkoliv například u zdroje Noaa1 by posun v řádu desítek minut mohl pomoci.

Výsledek po úpravě preprocessingu je zobrazen na obr. 16 a 17. Je evidentní, že došlo ke změně velikosti množiny dat. Přesto je možné alespoň částečně posoudit změnu MAE, protože do preprocessingu zdroje Hawaii nebylo mezi výstupy zasahováno. Důvodem pro zachování preprocessingu Hawaii je, že preprocessing byl upraven po odevzdání projektu II. Posun dat zdroje Hawaii je způsoben posuny kolem nulové polohy v jednotlivých měsících a nelze tak očekávat, že globální posun pomůže.

Úpravou došlo k zhoršení u zdroje Noaa1, jinak lze obecně říci, že data zbývajících zdrojů jsou nyní zpracována lépe.

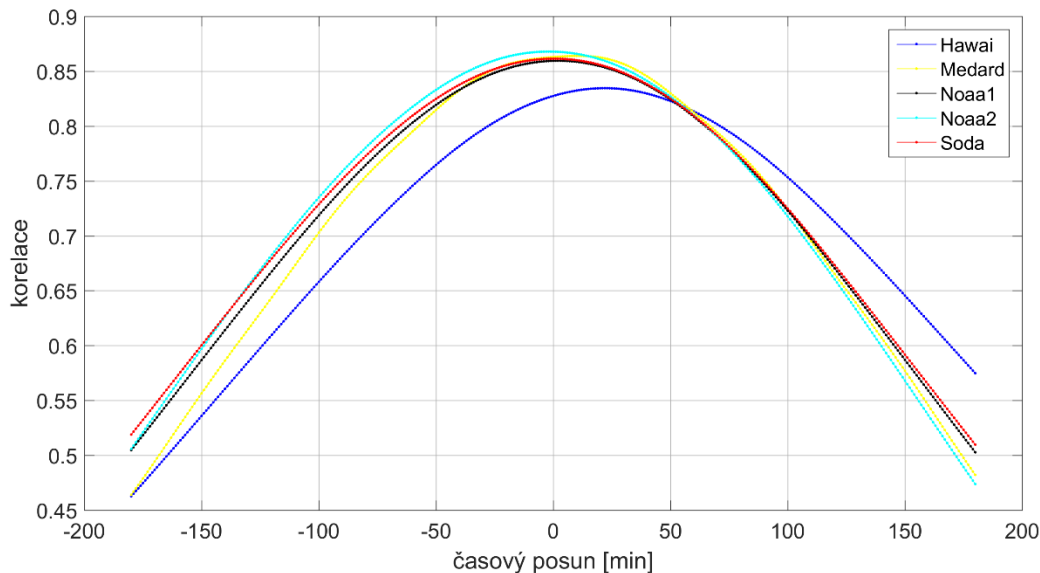


Obrázek 16: Vychýlení dat vůči realitě na celých datech (původní)



Obrázek 17: Vychýlení dat vůči realitě na celých datech (po úpravách)

Z pohledu korelace je posunování osvitů odpovídající, což je možné pozorovat na obr. 18. V tomto případě nebyl použit Spearmanův korelační koeficient, protože způsoboval nespojitosti na každé celé hodině. Pravděpodobně to souvisí s parazitními jevy při interpolaci během posunu dat.



Obrázek 18: Posun dat vůči realitě po úpravách, vypočtená korelace

## 7.2.2 Oblačnost

Zatímco u osvitů je vyrovnaní dat celkem jednoduchý úkol, protože je možné snadno porovnat data s naměřenými hodnotami. V případě oblačnosti je situace podstatně obtížnější, protože chybí naměřená data pro vyhodnocení posunů.

První řešení problému, které přicházelo v úvahu, byl posun podle korelace. Pro korelaci je nejprve potřeba data normalizovat, protože korelace dat o oblačnosti a osvitů je pochopitelně neúčinná a velmi malá. Byly vyzkoušeny 2 způsoby, jak data normalizovat:

- 1)  $CC_{norm} = (CC) * H$ ,
- 2)  $CC_{norm} = (CC) * G_{teor}$ ,

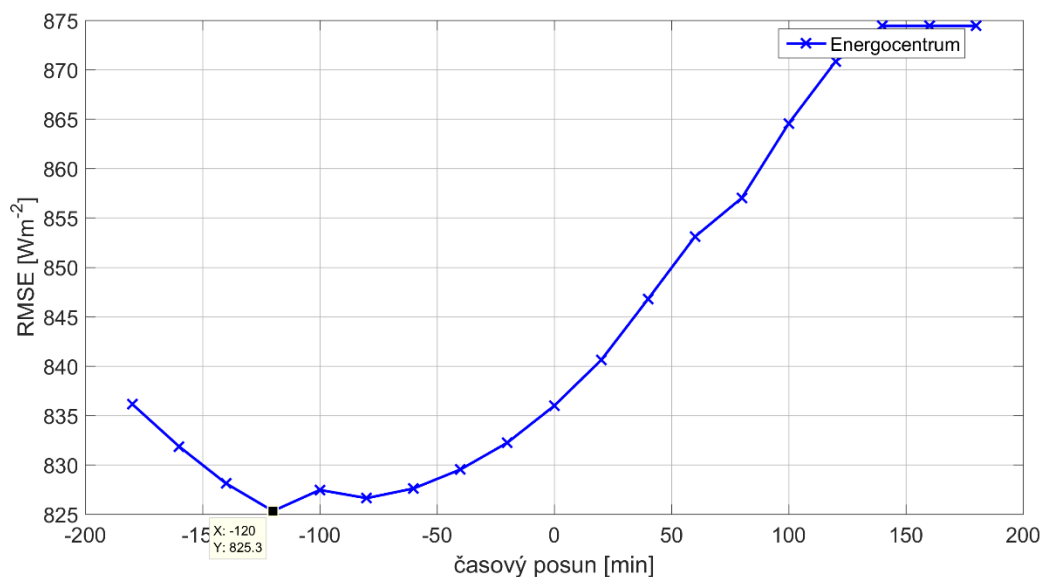
kde  $CC$  = oblačnost převedená na interval  $< 0; 1 >$ ,  $G_{teor}$  = teoretická hodnota osvitů pro konkrétní den a  $H$  = solární horizont. Noční hodiny jsou během normalizace nulovány.

Normalizace jako takové ale nejsou dostačující, protože pro každý zdroj je interpretace dat jiná, obvykle neznámá. Obvykle jde o číslo udávající množství mraků na obloze, útlum osvitů apod. I po četných pokusech bylo dosaženo jen nepřesvědčivých výstupů. Maximum se mezi jednotlivými pokusy obvykle posouvalo a bylo by obtížné určit, která normalizace je správná.

Výsledek navíc bylo potřeba posouvat podle osy  $y$ , protože maximální zakrytí neznamená, že výsledný osvit je nulový. Ke všemu se ještě objevovaly nespojitosti způsobené interpolací.

Dá se předpokládat, že jde o obecnou funkci, která může být, a pravděpodobně i je, nelineární a hledat její popis není z hlediska času levná úloha. Pokusy o normalizaci byly nakonec nahrazeny posunem, který se prováděl na vstupech neuronové sítě a následně se podle výstupů zjišťovala RMSE predikce. Grafy 13-16 ukazují výstupy neuronové sítě v závislosti na tom, jak se zakrytí oblohy posouvalo v čase.

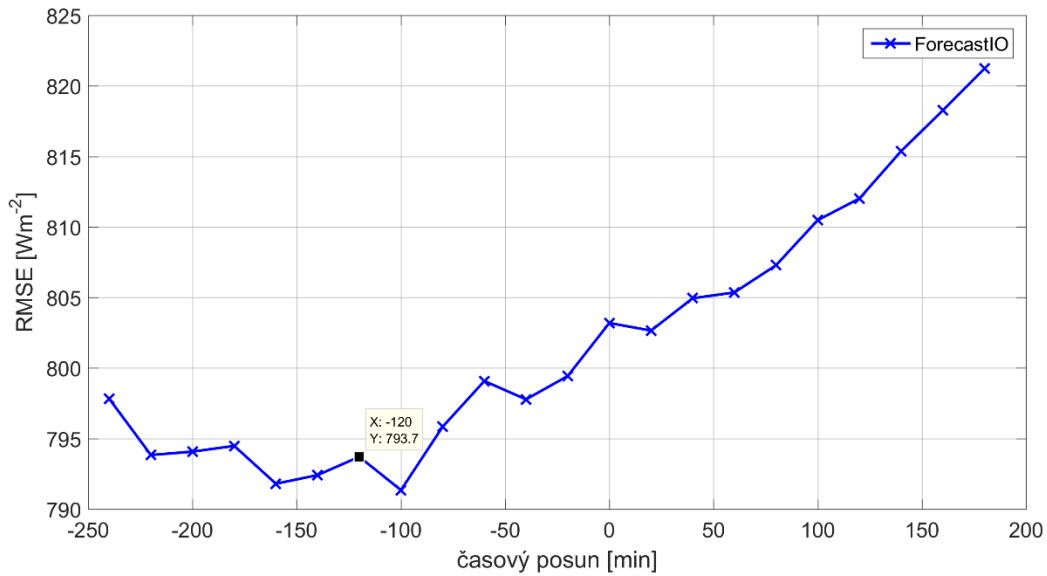
Nevýhodou této robustní metody je, že nutně neznamená, že data jsou po posunu přesnější, ale podobně jako v případě kombinování oblačnost s osvitem v podkapitole 8.4 *Závěrečné srovnání zdrojů* jsou nejvhodnější pro síť. Dokonce může nastat situace, kdy pro řešení úlohy neuronovou sítí je vhodnější na vstup dávat zakrytí oblohy úmyslně o hodinu posunuté (budoucí, nebo předchozí hodnotu). V této práci není řešeno, zda tomu tak je, ale při použití preprocessingu určeného pro tuto úlohu v jiných aplikacích je na místě na to brát ohled.



Obrázek 19: Posun dat oblačnosti pro zdroj Energocentrum, označený bod je časový posun, na který je preprocessing zdroje optimalizovaný

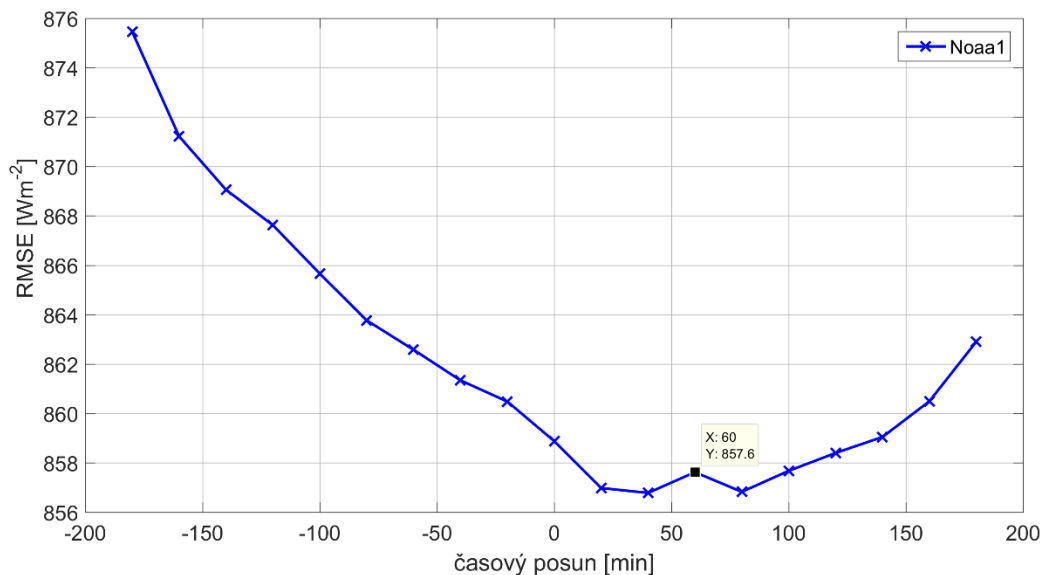
Pro energocentrum je patrné, že se vyplatí data posunout o 2 hodiny doleva po časové ose. Na první pohled je vidět, že graf obsahuje jasné globální minimum, což odpovídá všem předpokladům. Pro analýzu byla použita data od 1.2.2016 do 20.4.2017, některé vzorky chyběly, takže celkově se pracovalo s daty z 398 dní.





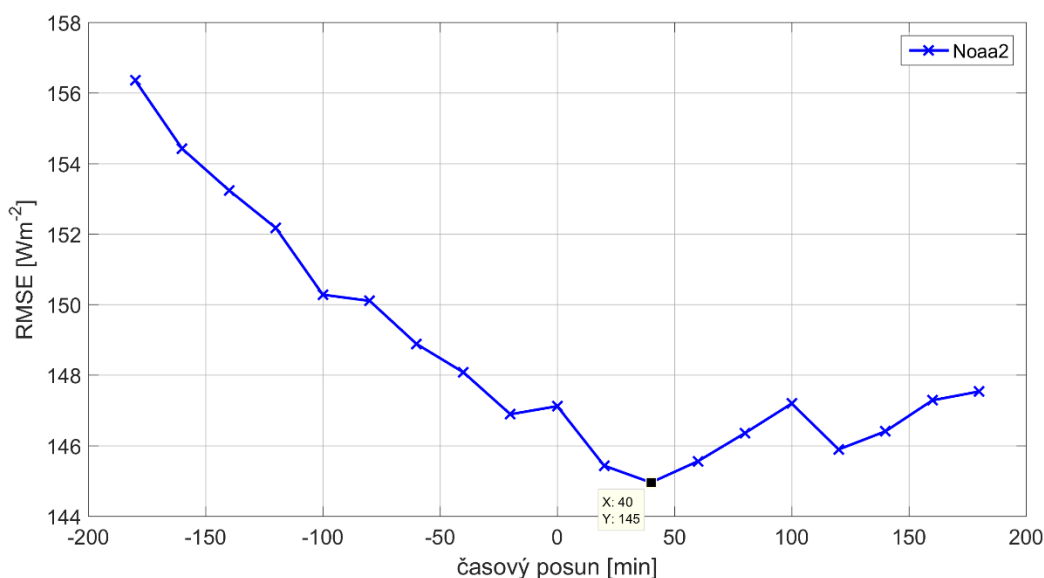
Obrázek 20: Posun dat oblačnosti pro zdroj Forecast IO, označený bod je časový posun, na který je preprocessing zdroje optimalizovaný

Stejně jako v případě energocentra jsou nakonec data posunuta o 2 hodiny. Je možné, že nejde o optimální posun, ale vzhledem k preprocessingu dat odpovídá rozdílu českého času oproti UTC. Bylo přihlédnuto k faktu, že rozdíl mezi globálním minimem a vybraným bodem je zanedbatelný a posunutí o minuty se obtížně vyhodnocuje vzhledem k nepřesnosti sítě při podmínkách nastavených pro výpočet této úlohy (více na konci kapitoly). Data byla zpracována v rozsahu od 11.2.2016 do 13.4.2017, opět nějaké vzorky chybí, tedy 343 dní.



Obrázek 21: Posun dat oblačnosti pro zdroj Noaa1, označený bod je časový posun, na který je preprocessing zdroje optimalizovaný

Data pro zdroj Noaa1 budou opět posunuta o hodinu, tentokrát opačným směrem. Noaa1 byla zpracována až na druhý pokus, neboť původní posun na datech byl tak velký, že se nevešel do šestihodinového okna, pro které byla úloha navržena. Ze současného stavu je nakonec patrné, že zobrazený graf má své globální minimum při posunutí někde v rozsahu 40-80 min, od odhadnuté hodnoty po prvním výpočtu, což stejně jako v případě Forecast IO bude zaokrouhлено na celé hodiny. Noaa1 je z hlediska výpadků patrně nejstabilnější zdroj, data byla vypracována od 10.2.2016 do 20.4.2016 a po zahrnutí výpadků má testovací množina 434 dní.



Obrázek 22: Posun dat oblačnosti pro zdroj Noaa2, označený bod je časový posun, na který je preprocessing zdroje optimalizovaný

Pro zdroj oblačnosti Noaa2 je vyhodnocení prakticky nemožné. Pravděpodobně se vyplatí data o 40 min posunout, ale vzhledem k tomu, že jde o zdroj dat, který se začal stahovat až kvůli této úloze a že učicí množina sítě je poměrně obsáhlá, je tento graf tvořen šestnáctidenním úsekem, což rozhodně není vhodné pro vyhodnocení celoročních trendů.

Tato úloha je zpracována poměrně povrchně a určitě je vhodné se k ní vrátit a zpracovat ji hlouběji. V novém zpracování bude možná vhodné grafy proložit křivkou a vyhodnotit minimum přesně, v tomto prvotním pokusu lze vystačit se zaokrouhlením na hodiny a prokládání křivkou postrádá smysl. Důvodem povrchního zpracování je časová náročnost úlohy v kombinaci s nedostatkem času. Toto zjednodušení je odpustitelné, protože z grafů je patrné, že zpřesnění posunů na výsledek bude mít minimální vliv. Aby byla úloha zpracována přesně je dobré zvýšit počet sítí pro predikci (např. na 1500-2500), v predikci na vyšší hodnotu a změnit vzorkování z 20 na 1-5 minut. Bohužel toto nastavení by i na poměrně výkonných

serverech (4x8 jader 2.4 Ghz, 8 GB RAM), které byly pro zpracování úlohy k dispozici trvalo několik týdnů.

V každém ohledu významným výstupem je znalost, kde minima leží, a je tedy možné prohledat o něco menší okolí.

### 7.3 Stanovení korelace dat dostupných zdrojů osvitů

Korelace vstupů a výstupů neuronové sítě dávají základní představu o informaci, kterou nesou. Je pochopitelné, že korelace budou poměrně vysoké už vzhledem k faktu, že noční hodnoty jsou vždy nulové a denní mají vždy pro průběh osvitů charakteristický tvar.

Pro výsledné korelace má v některých případech smysl noční hodiny uvažovat, protože u některých vstupů mají informační hodnotu. Například u osvitů je důležité, jestli Slunce začíná a přestává zářit ve správný čas. To je také dobře vidět na změně nejlepšího zdroje mezi sloupci v tabulce č.1.

Tabulka 1: Korelace predikcí osvitů po preprocessingu

Zdroj	$r_{sp0}$	$r_{sp-0}$	$r_{sp0}$ = Spearmanův korelační koeficient, noční hodnoty zahrnuté $r_{sp-0}$ = Spearmanův korelační koeficient, bez nočních hodnot.
Hawai	0,9573	0,8465	Jde o porovnání celých vektorů dat, ne o porovnávání po jednotlivých dnech. Pro vypracování tabulky byla použita data v rozmezí od 9.1.2016 do 30.3.2017.
Medard	0,9542	0,8871	
Noaa1	0,9437	0,8648	
Noaa2	0,9445	0,8577	
Soda	0,9440	0,8690	

V kapitole 10 *Zpracované výstupy předpovědi osvitů* budou uvedeny změny kvality sítě v závislosti na zvolení různých zdrojů vstupních dat při trénování sítě. To má pochopitelně vyšší vypovídající hodnotu, protože zahrnují všechny souvislosti přímo v rámci chování celého algoritmu.

### 7.4 Srovnání chyby sítě s různými zdroji osvitů

Tabulka 2: Srovnání zdrojů predikcí při stejných podmínkách

Zdroj	RMSE ( $Wm^{-2}$ )	RMSE <sub>rel</sub> (%)	Tabulka 2 zobrazuje výstupy sítě při použití různých dat. V tomto případě jde o věrohodnější posouzení kvality zpracovaných zdrojů dat než v případě korelací. Výstupy byly vypracovány pro data v časovém rozmezí od 4.1.2016 do 29.3.2017.
Hawai	828	23,7	
Medard	816	23,4	
Noaa1	835	23,9	
Noaa2	833	24,0	
Soda	852	24,3	

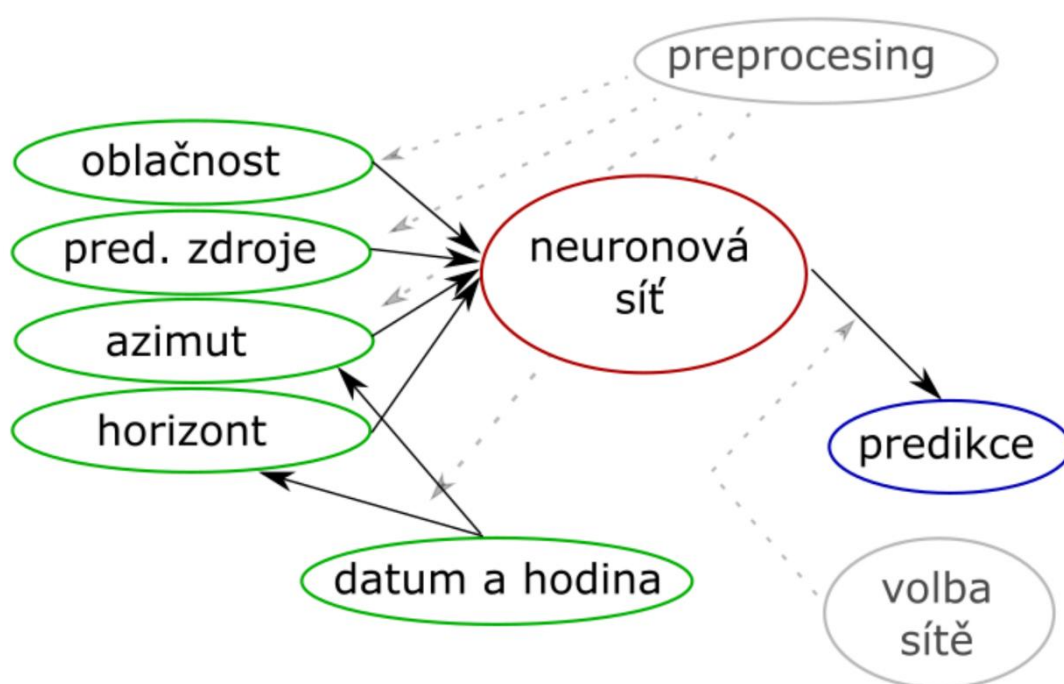
Je patrné, že zdroj Medard se ukázal jako nejpresnější. Bohužel to, zda bude nakonec použit není jen otázka kvality, ale také dostupnosti pro komerční účely a zdroj Medard je placený.

Předběžně se tedy zdá, že nejvhodnějším zdrojem je zdroj Hawai, který jak už bylo zmíněno výše je hlavním zdrojem stávajícího algoritmu. Zdroje Noaa1 a Noaa2 jsou na tom podobně, a mohou být bezpochyby použity jako náhradní zdroje v případě výpadku, ovšem u Noaa1 bude pro tento případ na stránkách potřeba uvést, že je použit. Pro finální podobu algoritmu bude ale podstatné, jak funguje interakce všech vstupních parametrů se sítí. To bude zvaženo v podkapitole 8.4 *Závěrečné srovnání zdrojů* po odladění parametrů sítě.

Zdroj dat, který byl určen jako nejpřesnější je nejpřesnější vzhledem k sítí, ale obecně to platit nemusí. Síť je schopna naučit se některé tendence dat. Pokud bude například zdroj špatný, ale vždy stejným způsobem, například bude podávat 1.5x větší hodnotu, než je hodnota očekávaná, síť je schopná se s tímto problémem vyrovnat. To může spolu s dalšími faktory přispívat k jinému výsledku než při výpočtu korelací v předchozí podkapitole.

## 8 Aktuální síť

V současné době se pracuje s poměrně jednoduchou sítí viz (obr. 23), která využívá znalosti datumu a hodiny, na kterou se předpovídá, jako pevné vstupní hodnoty ve snaze dosáhnout co nejvyšší stability předložením prokazatelně správných dat. Na rozdíl od těchto dvou vstupů je zakrytí oblohy a predikce z jiného zdroje zpravidla zatížena chybou. Výstupní predikce je pro proces učení nahrazena naměřenými daty, která jsou pochopitelně rovněž zatížena chybou, ale ve srovnání s ostatními vlivy je tato až na snadno detekovatelné výjimky (např. úplně vypnutá čidla) zcela zanedbatelná.



Obrázek 23: Neuronová síť, stávající model

### 8.1 Trénování neuronové sítě

Samotné trénování neuronové sítě probíhá ve funkci `createAndTrainNetwork.m`. Pro trénování byl na základě dokumentace nástroje MATLAB [25] zvolen algoritmus `trainlm.m`, který byl rovněž zmíněn jako vhodný v nezávislém článku [26] (kapitola 3.4).

Neuronová síť, která se učí na nepřesných datech, má velmi mnoho stupňů volnosti, které souvisí s prostorem chyb určovaných při trénování sítě. Algoritmy, pro trénování sítě, mezi které patří i Levenberg-Marquardt backpropagation, jsou založeny na prohledávání prostoru vah sítě. V případě Levenberg-Marquardt backpropagation se k prohledávání používá

kombinace gradientní metody a metody nejmenších čtverců. Jedná se zde o gradient chyby sítě. Zjednodušeně si ho lze představit jako gradient vícerozměrné funkce, která z množiny vah všech neuronů vypočte chybu sítě. To, že síť má mnoho stupňů volnosti pak znamená, že kromě globálních minim (hledá se minimální chyba) je v prostoru velmi vysoký počet minim lokálních, do kterých má síť při tréninku tendence sklouznout. Proto je pro předpověď připraven větší počet sítí, určený vstupním parametrem<sup>vi</sup> příslušné funkce (viz 8.3 *Funkce pro vytvoření a trénování sítě*).

Původní myšlenka byla z množiny natrénovaných sítí se vybrat tu „nejlepší“. Otázkou je, jak takovou síť vybrat.

První cestou, která se zde nabízí, je stanovení rozumné kriteriální funkce a na základě její hodnoty volit optimální síť. Během řešení tohoto úkolu bylo vyzkoušeno vícero kriteriálních funkcí. Většina z nich byla poměrně jednoduchá, jako např. průměr RMSE chyby v několika předchozích dnech apod.

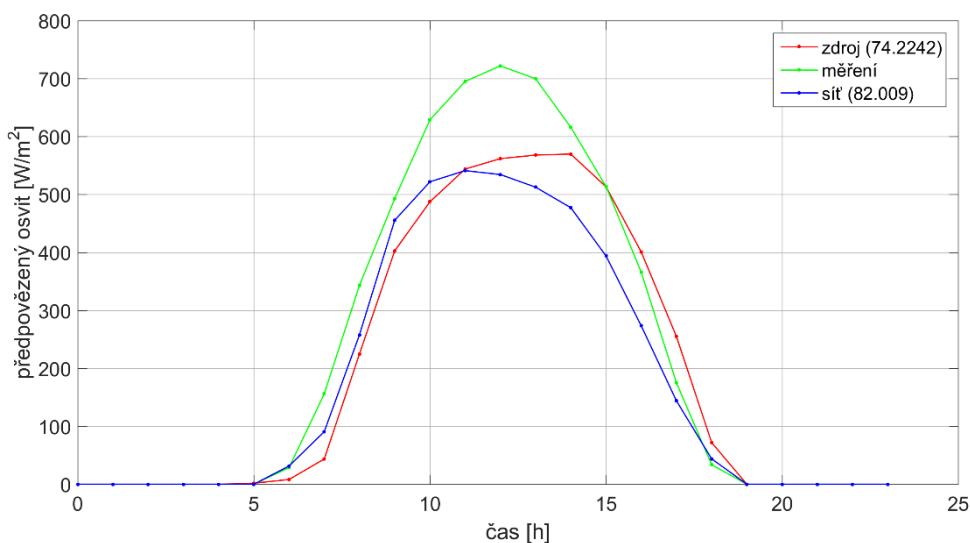
Za předpokladu, že funkce užívá například chybu MAE, může nastat následující problém. Tato chyba zohledňuje vyšší hodnoty více než malé. Pokud algoritmus předpovídá na den, kdy bude výsledný osvit  $300 \text{ Wm}^{-2}$  a někdy v předchozích dnech bylo  $400\text{-}800 \text{ Wm}^{-2}$ , dojde k tomu, že velké hodnoty budou mít na výběr větší vliv než malé. Proto pro dny kolem  $800 \text{ Wm}^{-2}$  budou mít největší váhu, a získá se nevhodná síť, která je lépe natrénována na hodnoty vysoko nad oněch  $300 \text{ Wm}^{-2}$ . Relativní chyby zase mají obdobný problém s malými hodnotami. V případě použití RMSE je vliv velkých chyb, které odpovídají vyšším hodnotám ještě zesílen.

Další možnost, je algoritmu zkusit podstrčit nějakou další informaci, která má tento problém eliminovat. Typické řešení je třeba vybírat jen dny, které jsou nejpodobnější. Jedním z důvodů, proč se tato cesta zdá nevhodnou, je to, že jde v podstatě o snahu podstrčit síti informaci, která už byla nějakým způsobem zahrnuta. V algoritmu už je zahrnuta část, kde se veškeré informace, které k tomuto účelu lze použít, už jednou využily. Data jsou navíc poměrně nepřesná a experimentální metodou byly často nalezeny sítě, kdy pro dva stejné vstupy ze zdrojové predikce byly naměřeny dva úplně různé průběhy (obr. 24 a 25). Z toho plyne,

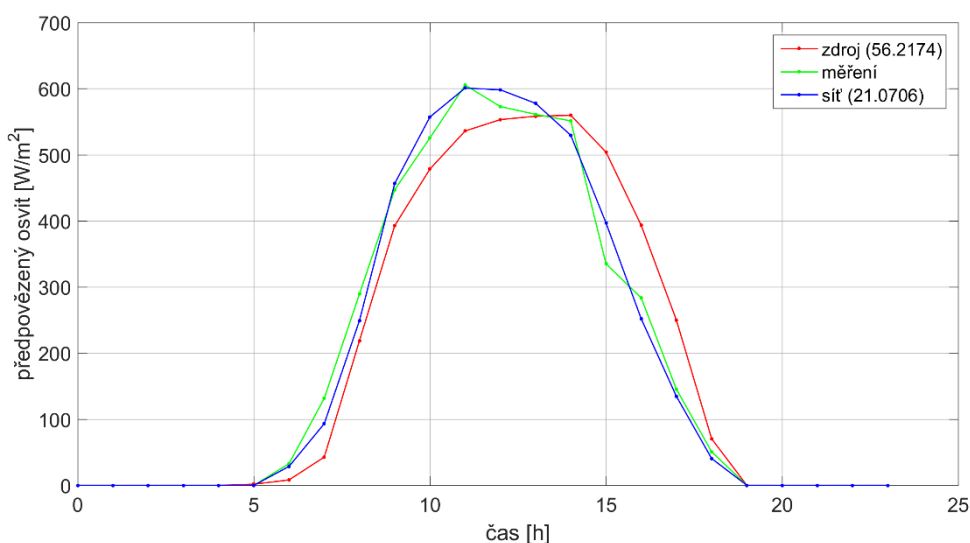
---

<sup>vi</sup> Obvykle 500-1000.

že i když se podaří vybrat nejlepší síť pro z nějakého hlediska nejpodobnější dny, problém částečně přetrvává.



Obrázek 24: Srovnání (2.4.2016), při stejné predikci zdroje jsou změřeným výstupem 2 odlišné průběhy viz obr. 25



Obrázek 25: Srovnání (3.4.2016), při stejné predikci zdroje jsou změřeným výstupem 2 odlišné průběhy viz obr. 24

Nakonec se jako nejvhodnější ukázalo volbu sítě vůbec neprovádět a předpovídat predikci pomocí náhodné množiny sítí. Už v projektu I [1] bylo patrné, že ačkoliv některé sítě nejsou natrénovány dobře, průměrně natrénovaná síť je lepší než PV Forecast. Původním pohledem na problematiku tedy bylo snažit se vybrat jednu síť, která bude nejlepší podle nějakého kritéria. Nově nalezeným optimálním pohledem je nechat hádat náhodnou množinu sítí.

Abychom demonstrovali, že zvolená metoda dosahuje ve srovnání s jinými dobrých výsledků, byla provedena predikce vícero metodami. Data pro predikci vyhodnocenou v tabulkách 3 a 4

byla v rozsahu 1.4.2016 až 28.2.2015. Směrodatná odchylka (RMSE) výstupu byla počítána pro vektory dat z celého měsíce. Vlivem chyby ve skriptu byla v jednom běhu při výběru uvažována jen natrénovaná síť, která byla nejbližší naměřeným datům, a to pro celý rok. Výsledná hodinová RMSE této predikce (v tabulkách označená jako Ideal) klesla o více než 30 % oproti hodnotě, kdy byla síť vybírána pomocí kriteriální funkce, a síť pochopitelně předpovídala velmi dobře. Je to zajímavá teoretická hodnota, které sice z principu nelze vhodným výběrem dosáhnout, ale ukazuje, jak daleko je neoptimalnější dosažitelná hodnota. Výstup této metody je ve zmiňovaných tabulkách označen červeně.

Následuje stručný soupis významů zkratk v tabulkách 3 a 4:

- Democratic: Nesnaží se volit podobné dny, nebo je nějak hodnotit, jen natrénuje vyšší počet sítí a výsledná předpověď je průměr jednotlivých předpovědí.
- Onerun: Je předpověď jednou sítí. Síť není vybírána, prostě se připraví a natrénuje jen jeden exemplář, který provede celou předpověď.

Jména dalších použitých metod jsou složena ze slov, která popisují, jak je vhodná síť volena:

- MAErel: Síť se vybírá pomocí  $MAE_{rel}$  na pevně stanovených dnech, jde o každý druhý den za posledních 10 dní od data, na které je předpovídáno.
- MAErelelitistic: Funguje stejně jako  $MAE_{rel}$ , jen konečná předpověď je průměr lepší poloviny nalezených sítí.
- MAErelrealtrapz: Stejně jako  $MAE_{rel}$ , ale chyba není počítána na každém druhém z posledních deseti dní, ale na deseti nejpodobnějších z posledních třiceti dní. Podobnost se hodnotí jako  $MAE_{rel}$  mezi denní sumou předpovědi zdroje a skutečně naměřených hodnot posledních třiceti dní.
- MAErelrealtrapzelitistic: Stejně jako „MAErelrealtrapz“, s tou výjimkou, že konečná předpověď je průměr lepší poloviny nalezených sítí.
- MAErelsourcetrappz: Stejně jako „MAErelrealtrapz“, ale podobnost je mezi předpovědí zdroje ze dne, na který se předpovídá a předpověďmi výchozím zdrojem, které byly během posledních třicet dní.
- MAErelsourcetrappzelitistic: Stejně jako „MAErelsourcetrappz“, jen konečná předpověď je průměr lepší poloviny nalezených sítí.

Analogicky „corrrealelitistic“ aj. fungují na stejném principu jako „MAErelelitistic“, jen místo MAE se počítá korelace.



Z tabulek 3 a 4 lze potvrdit domněnku, že nejvhodnější je použít pro výběr sítě metodu „democratic“, která vlastně síť vůbec nevybírání, ale použije všechny sítě bez ohledu na jejich naučení. Je patrné, že sítě s klíčovým slovem „elitistic“, které fungují stejně, ale snaží se vybrat ty nejlepší sítě, než jsou použity pro společnou předpověď dosahují víceméně stejných hodnot při početně náročnějším postupu.

Tabulka 3: Porovnání výběrů vhodné sítě (RMSE suma z jednotlivých měsíců), červeně označený řádek je teoreticky nejlepší možný výběr, zelený řádek je algoritmus, který je implementován

Použitá metoda	pokus 1	pokus 2	pokus 3	pokus 4	pokus 5	pokus 6	pokus 7	pokus 8
MAErel	848,69	839,39	855,69	868,68	952,00	830,87	852,68	847,09
MAErelelitistic	793,04	789,83	790,18	790,76	791,55	790,63	790,43	791,86
MAErelrealtrapz	802,44	816,41	784,10	813,17	844,32	834,65	866,21	808,14
MAErelrealtrapzelitistic	790,65	791,18	793,59	792,97	790,32	791,08	791,59	790,87
MAErelsourcetrappz	856,33	858,00	873,61	880,06	973,79	849,89	856,35	857,87
MAErelsourcetrappzelitistic	790,11	788,96	792,03	789,00	790,32	789,28	789,72	791,74
RMSE	855,39	838,21	871,18	873,40	860,03	871,79	857,14	844,13
RMSEelitistic	790,15	790,31	792,36	791,17	791,08	791,37	790,02	791,70
RMSErealtrapz	774,09	784,44	801,09	784,27	883,91	852,66	788,17	821,31
RMSErealtrapzelitistic	791,35	791,25	790,42	792,40	793,93	789,93	790,61	789,77
RMSEsourcetrappz	867,77	900,11	869,40	841,06	845,72	860,50	884,41	870,72
RMSEsourcetrappzelitistic	789,96	790,08	789,64	790,67	789,66	788,75	789,60	789,74
Corr	828,24	830,14	826,76	834,07	816,78	821,75	830,13	834,23
Corrrealelitistic	791,27	789,53	790,48	789,64	792,17	791,52	789,86	789,06
Corrrealtrapz	834,42	820,73	808,18	815,52	818,29	825,05	817,27	823,58
Corrrealtrapzelitistic	788,85	789,60	786,38	790,77	790,49	789,80	793,82	789,07
Corrsourcetrappz	830,95	820,52	823,31	851,64	819,77	838,32	826,61	824,19
Corrsourcetrappzelitistic	790,30	790,64	790,34	790,81	789,49	790,65	790,06	789,74
Democratic	790,33	790,26	790,41	789,64	791,20	791,23	789,46	794,97
Ideal	525,67	531,17	525,95	528,81	524,66	525,80	525,25	522,29
Onerun	830,99	822,98	830,97	810,49	835,19	823,29	817,98	813,93
Trapz	906,83	851,84	848,98	950,54	847,49	880,33	862,63	863,25
Trapzelitistic	791,25	791,06	790,53	791,04	790,94	790,29	789,61	791,41
Trapzrealtrapz	826,29	789,81	780,13	787,55	798,43	771,22	787,41	757,23
Trapzrealtrapzelitistic	790,30	790,76	791,04	791,86	791,64	788,78	790,32	789,80
Trapzsourcetrappz	843,13	854,68	876,57	869,44	866,51	847,59	845,93	854,91
Trapzsourcetrappzelitistic	794,29	789,84	790,65	793,15	790,94	793,21	790,57	790,14

Za zmínku stojí i metoda „RMSErealtrapz“, která v prvním pokusu dosáhla ještě podstatně lepšího výsledku. Vezme-li se ale v úvahu dalších 7 pokusů, je zřejmé, že jde o nestabilní, a tedy nevhodný algoritmus. Rozsah, ve kterém se výstupy pohybují, je dokonce ještě širší než v případě, že se síť vůbec nevolí („onerun“).

Tabulka 4: Porovnání výběrů vhodné sítě (RMSErel suma z jednotlivých měsíců), červeně označený řádek je teoreticky nejlepší možný výběr, zelený řádek je algoritmus, který je implementován

Použitá metoda	pokus 1	pokus 2	pokus 3	pokus 4	pokus 5	pokus 6	pokus 7	pokus 8
MAErel	0,26	0,26	0,27	0,27	0,30	0,26	0,27	0,26
MAErelitistic	0,25	0,24	0,24	0,25	0,25	0,24	0,24	0,25
MAErealtrapz	0,25	0,25	0,24	0,25	0,27	0,26	0,27	0,25
MAErealtrapzelitistic	0,24	0,24	0,25	0,25	0,24	0,24	0,25	0,24
MAErelsourcetrappz	0,27	0,27	0,27	0,28	0,32	0,27	0,27	0,27
MAErelsourcetrappzelitistic	0,24	0,24	0,25	0,24	0,24	0,24	0,24	0,25
RMSE	0,26	0,26	0,27	0,27	0,26	0,27	0,26	0,26
RMSEelitistic	0,24	0,24	0,25	0,25	0,25	0,24	0,24	0,25
RMSErealtrapz	0,24	0,24	0,25	0,24	0,27	0,27	0,24	0,25
RMSErealtrapzelitistic	0,25	0,25	0,24	0,25	0,25	0,24	0,24	0,24
RMSEsourcetrappz	0,27	0,28	0,27	0,26	0,26	0,27	0,28	0,27
RMSEsourcetrappzelitistic	0,24	0,24	0,24	0,24	0,24	0,24	0,24	0,24
Corr	0,26	0,26	0,26	0,26	0,26	0,26	0,26	0,26
correalitistic	0,25	0,24	0,24	0,24	0,25	0,24	0,24	0,24
correaltrapz	0,26	0,26	0,25	0,26	0,26	0,26	0,26	0,26
correaltrapzelitistic	0,24	0,24	0,24	0,24	0,24	0,24	0,25	0,24
corrsourcetrappz	0,26	0,26	0,26	0,27	0,26	0,26	0,26	0,26
corrsourcetrappzelitistic	0,24	0,24	0,24	0,24	0,24	0,24	0,24	0,24
democratic	0,24	0,24	0,24	0,24	0,24	0,25	0,24	0,25
ideal	0,15	0,16	0,15	0,15	0,15	0,15	0,15	0,15
onerun	0,26	0,26	0,26	0,25	0,26	0,26	0,25	0,25
trapz	0,27	0,26	0,26	0,29	0,26	0,27	0,26	0,26
trapzelitistic	0,25	0,25	0,24	0,25	0,24	0,24	0,24	0,25
trapzrealtrapz	0,26	0,24	0,24	0,25	0,25	0,24	0,24	0,23
trapzrealtrapzelitistic	0,24	0,24	0,24	0,25	0,25	0,24	0,24	0,24
trapzsourcetrappz	0,26	0,26	0,27	0,27	0,27	0,26	0,26	0,26
trapzsourcetrappzelitistic	0,25	0,24	0,24	0,25	0,24	0,25	0,25	0,24

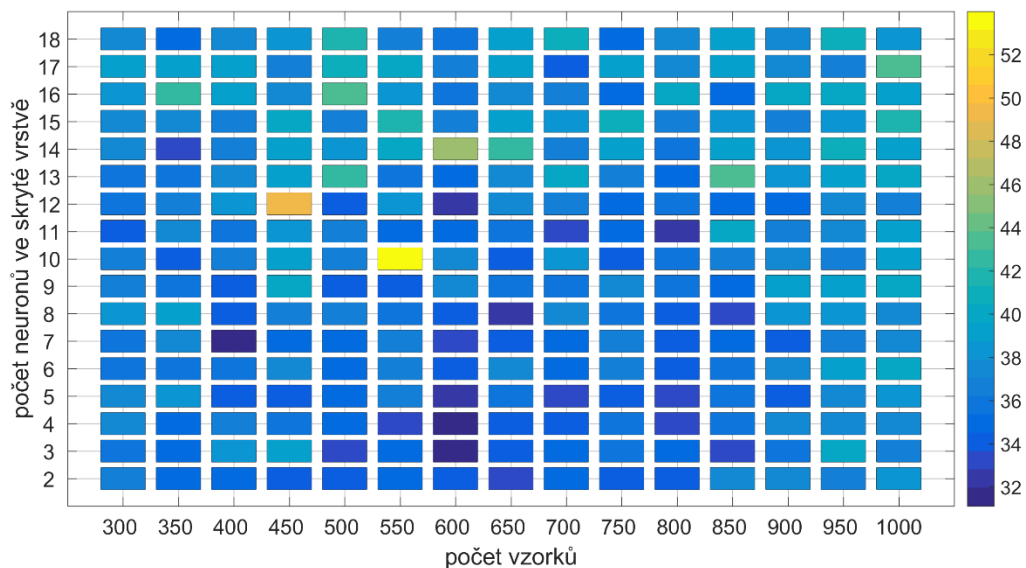
V samotném závěru trénovacího algoritmu je ještě zařazena jednoduchá analýza výstupu, která má za úkol odstranit velmi chybné předpovědi. Například v případě, že síť předpoví zápornou hodnotu během dne, je predikce pro tento den přepočítávána, dokud nedojde k nezáporným hodnotám (dříve jen nulována). Obvykle by tato situace nastat neměla, je to spíše pojistka. Tento problém obvykle nastává při malém počtu cyklů nebo při použití některých méně robustních algoritmů pro výběr neuronové sítě z množiny.

## 8.2 Hledání vhodné topologie sítě

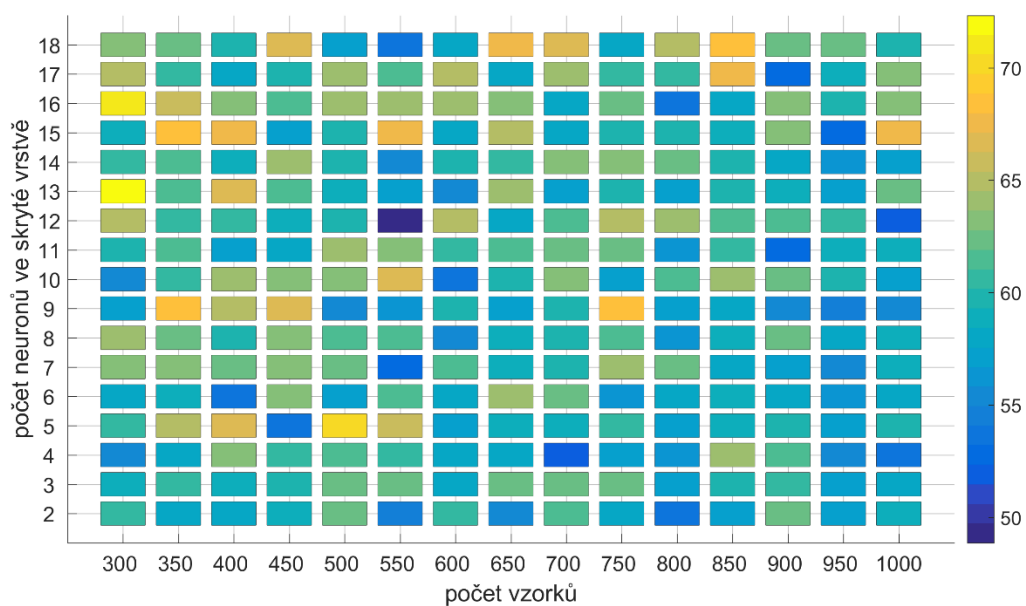
V počátcích práce, tzn. během projektu I [1] a projektu II [2], se obvykle pracovalo se sítí o konstantním počtu neuronů ve skryté vrstvě a konstantním počtem vzorků pro učení. Vzhledem k tomu, že výstupem této práce má být odladěný algoritmus, je třeba z tohoto hlediska síť přehodnotit.

Vzhledem k tomu, že jde o početně náročnou úlohu, byl sehnán poměrně výkonný server a byly propočítány chyby predikcí sítí v době od 1.4.2016 do března 2017. Počet neuronů ve skryté vrstvě byl v rozsahu 2-18, počet vzorků vstupních dat od 300 do 1000. Vše bylo spuštěno pro zdroj Hawaii pro osvit, Forecast IO doplněný v případech výpadku o Energocentrum pro oblačnost.

Výstupy tohoto postupu jsou interpretovány v podobě grafů jako jsou ty na obr. 26, 27, 28 a 29.



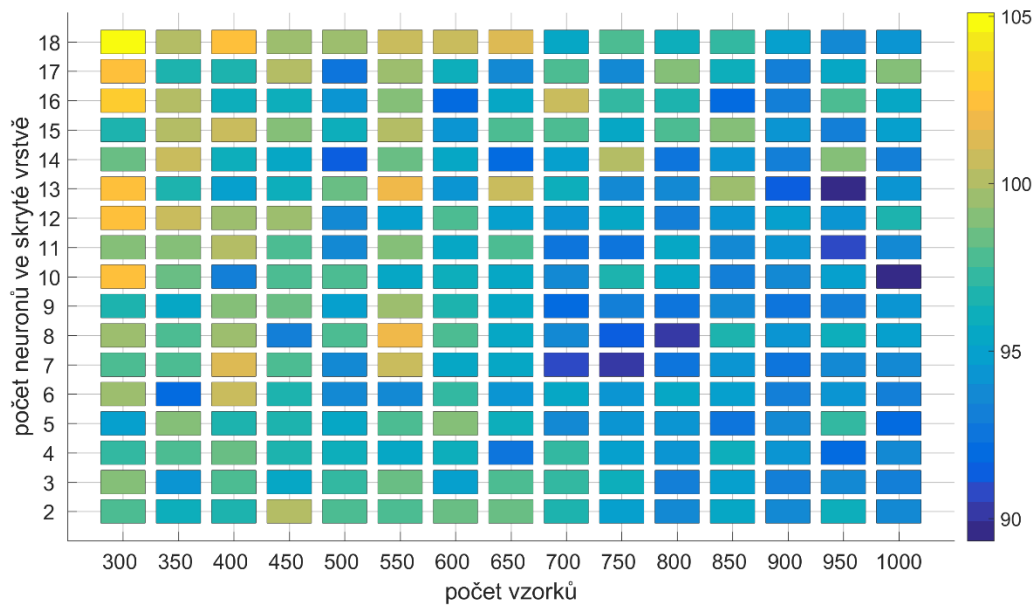
Obrázek 26: Příklad 1, výstupní graf z úlohy serveru pro interval 26.1.2017 – 2.10.2017



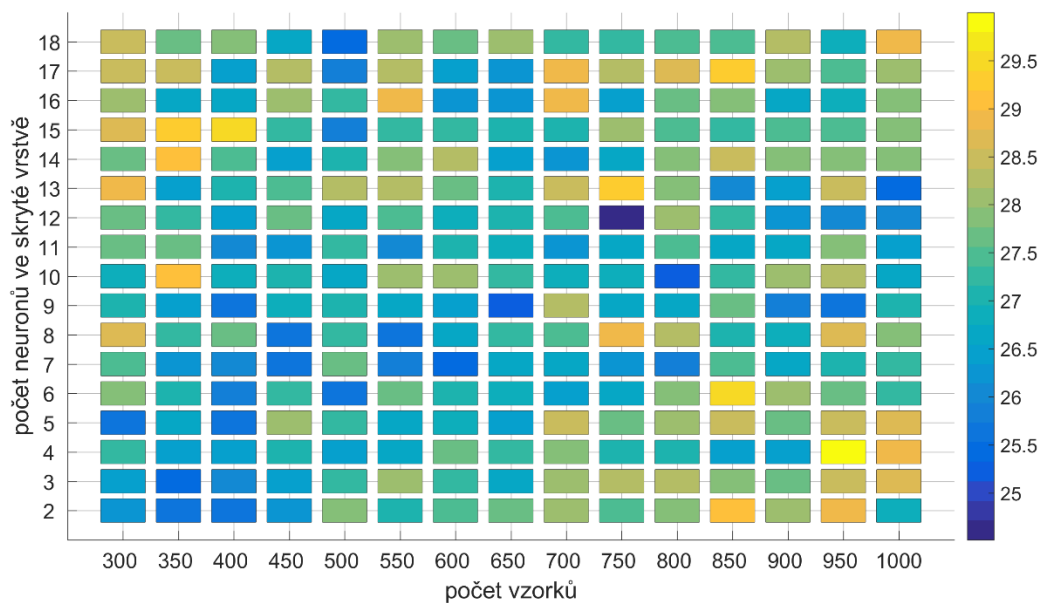
Obrázek 27: Příklad 2, výstupní graf z úlohy serveru pro interval 2.10.2017-25.2.2017

Z grafu na obr. 26 a 27 je patrné, že i pro 2 po sobě jdoucí krátké časové úseky dochází k posunu globálního minima. Je proto obtížné vyvodit z výstupů pevné výsledky.

Obecně lze pozorovat, že počet vzorků je závislý na délkách dní v daném období, což dává smysl vzhledem k tomu, že pro trénování sítě se uvažují jen nenoční hodnoty. To je dobře vidět na rozdílu mezi grafy na obr. 28 a 29.



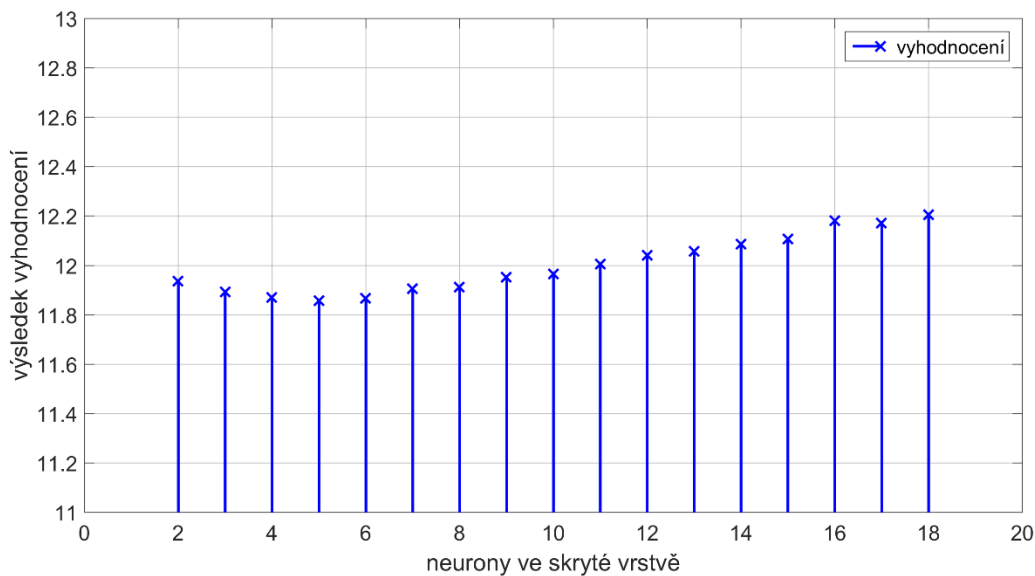
Obrázek 28: Příklad 3, výstupní graf z úlohy serveru pro interval 30.7.2016-29.8.2016



Obrázek 29: Příklad 4, výstupní graf z úlohy serveru pro interval 27.11.2016-27.12.2016

Co se týká složitosti sítě, je z prožití všech grafů (všechny jsou přiloženy v příslušné úloze na paměťovém médiu) patrné, že závisí nejen na konkrétním časovém úseku v roce (např. měsíci), ale i na konkrétním dni. Vzhledem k tomu, že vyhodnocení této závislosti není jednoduše zvládnutelná úloha, byl na základě výstupů v dalším odstavci popsané úlohy, počet neuronů ve skryté vrstvě, stanoven na 5.

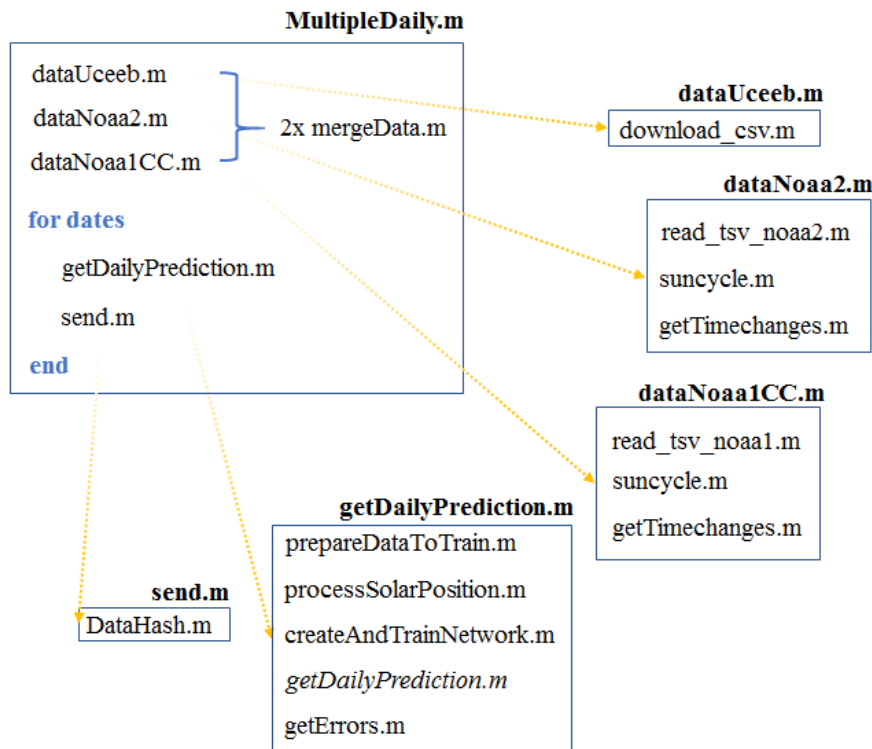
Pro nalezení vhodného počtu neuronů byl nejprve vyhodnocen počet vzorků pro jednotlivé měsíce v závislosti na předchozí úloze. Dále byly v rozsahu jednoho roku (1.3.2016 až 28.2.2017) propočítány sítě o 2 až 18 neuronech. To vše je provedeno 2x pro potvrzení výsledku. Aby se nevyskytl již zmiňovaný problém, že letní hodnoty by pro chybu RMSE opět měly větší váhu než zimní, byly chyby RMSE určeny pro každý měsíc zvlášť, dále jsou převedeny na chyby relativní a to tak, že pro jeden každý měsíc se chyba z konkrétní sítě vztahovala průměrné chybě sítě v daném měsíci a poté byly sečteny. Soudě dle grafu na obr. 30 je nejvhodnější síť tímto postupem dohledatelná a jasně ukazuje na výše zmíněný počet neuronů. Také lze usoudit, že od určité hodnoty už nemá smysl zkoušet další, složitější síť, protože chyba s rostoucím počtem neuronů od určité hodnoty poměrně prudce narůstá.



Obrázek 30: Hledání vhodného počtu neuronů pro předpověď osvitů ČVUT UCEEB

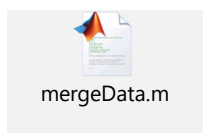
### 8.3 Funkce pro vytvoření a trénování sítě

V této podkapitole je sestaven stručný popis funkcí a skriptů, ve kterých je nový algoritmus naimplementován. Na obr. 31 je zobrazeno schéma celého algoritmu.



Obrázek 31: Schéma algoritmu na principu neuronových sítí

#### 8.3.1 Dodatečný preprocessing dat před vstupem do sítě



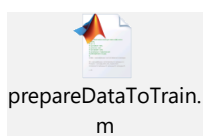
mergeData.m

Funkce spojuje matice dat na vstupu do jednoho výstupu. V algoritmu ze dvou vstupních matic  $M^{n,5}$  udělá jednu výstupní matici ve formátu  $M^{n,6}$  tak, že podle časové značky v prvních čtyřech sloupcích matic spojí poslední sloupec (sloupec s daty) do jedné matice. Přitom vyřadí nepřřazené vzorky dat (celé řádky).



processSolarPosition.m

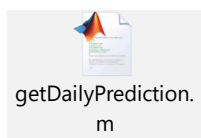
Slouží k přepočítání časových značek na polohu slunce. Funkce rovněž vyřadí hodnoty, kdy je slunce schované za horizontem v závislosti na vstupním parametru, a vrátí vektor složený z prvků  $\{0,1\}$ , kde 0 znamená, že hodnota na této pozici byla vyřazena a 1 naopak.



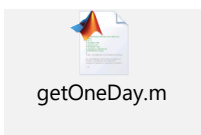
prepareDataToTrain.m

Slouží k vybrání dat pro trénování sítě. Náhodně promíchá data a rozdělí je na vstupy a známé výstupy sítě. Rovněž transponuje matici dat na učení tak, jak vyžaduje funkce, která slouží k trénování.

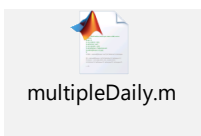
### 8.3.2 Skripty pro vytvoření a trénink sítě



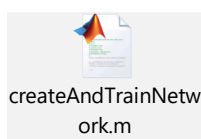
Předpoví osvit pro jeden den podle konkrétního data. Vstupem do funkce je i počet vzorků z historie, které má funkce pro trénink sítě uvažovat a počet sítí, ze kterých vybere jednu nejlepší podle RMSE. K tomuto vyhodnocení je použita funkce *getOneDay.m*. Výstupem funkce je proměnná *irradiation*, která obsahuje vstupy a nasimulované výstupy „nejlépe“ natrénované sítě.



Síť počítá predikci na jeden den podle data. Vstupem je na rozdíl od funkce *getDailyPrediction.m* neuronová síť, která je k tomuto účelu použita. Výstupem funkce je proměnná *irradiation*, která obsahuje vstupy a nasimulované výstupy sítě (pro stávající metodu výběru sítě není potřeba).



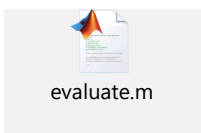
Skript, který předpoví osvit pro dny od počátečního do koncového data, pokud je to možné (data mohou chybět). Na začátku skriptu se nahrávají a upravují data do očekávané formy (viz 7 *Navržená řešení pro preprocesing dat*). Zastřešuje celý algoritmus, tj. pro seběhnutí algoritmu není potřeba nic jiného než pustit tento skript.



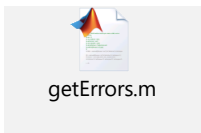
Vytvoří neuronovou síť a natrénuje ji, výstupem je natrénovaná neuronová síť.

### 8.3.3 Funkce pro různá zpracování výsledků

Celou práci provází různá vyhodnocování a zpracování výsledků. Ke konci práce byly úlohy pro přehlednost ukládány do navzájem si podobných struktur. Jde o úlohy ve složkách označených jako „computer task“ nebo „server task“ s popiskem úlohy. Tyto úlohy jsou



vyhodnocovány skripty *evaluate.m*, které se od sebe liší podle konkrétních požadavků. Skripty vždy pracují s daty v adresáři *storage\*, ve kterém jsou uloženy, stejně jako výsledky exekuce dané úlohy.



Funkce, která vrací chyby vypočtené mezi dvěma vektory podle prvních dvou bodů kapitoly 6 *Aparát užitý pro výpočty chyb* (Nezahrnuje korelace).

## 8.4 Závěrečné srovnání zdrojů

Pro srovnání zdrojů je důležité, aby byly vzájemně srovnávány predikce na stejná data. To je poměrně složité, protože u některých zdrojů jsou úseky, kde data prostě chybí. Situace se řeší tak, že je puštěn algoritmus s různými kombinacemi pro určité časové období a chybějící

místa se zaplní zdrojem Noaa1, který má chybějících dat nejméně. V případě Noaa1 se chybějící data ignorují. To má zaručit, že bude objeven nejlepší zdroj, nebo kombinace zdrojů, což ovšem nevadí, protože pro neuronovou síť bude stejně potřeba případné výpadky nahradit jiným zdrojem.

Data pro všechny zdroje s výjimkou oblačnosti Noaa2, která bude rozebrán zvlášť, jsou z časového rozmezí od 17.2.2016 do 29.3.2017. Srovnání jednotlivých zdrojů ukazuje následující tabulka:

Tabulka 5: Srovnání zdrojů, červeně označená je nejlepší hodnota, zeleně označená je vybraná hodnota

RMSE		zdroj osvitu				
		Hawai	Medard	Noaa1	Noaa2	Soda
zdroj oblačnosti	Eng	805,8	808,7	805,2	781,3	809,3
	IO	835,5	821,5	835,0	836,5	856,4
	Noaa1	832,9	812,3	842,6	818,2	848,7

RMSE daily sums		zdroj osvitu				
		Hawai	Medard	Noaa1	Noaa2	Soda
zdroj oblačnosti	Eng	8175	8267	8225	7639	8162
	IO	8654	8368	8709	8573	8991
	Noaa1	8691	8284	9017	8300	9012

RMSE <sub>rel</sub>		zdroj osvitu				
		Hawai	Medard	Noaa1	Noaa2	Soda
zdroj oblačnosti	Eng	2,7	2,7	2,7	2,6	2,7
	IO	2,8	2,7	2,8	2,8	2,8
	Noaa1	2,8	2,7	2,8	2,7	2,8

RMSE <sub>rel</sub> daily sums		zdroj osvitu				
		Hawai	Medard	Noaa1	Noaa2	Soda
zdroj oblačnosti	Eng	27,8	28,3	28,0	26,4	27,9
	IO	28,5	27,8	28,7	28,5	29,6
	Noaa1	29,1	28,1	30,0	28,1	30,2

Je patrné, že bez ohledu, zda jde o denní sumy nebo hodinová data jsou výsledky srovnatelné. Bezkonkurenčně nejlépe dopadlo spojení zdroje osvitu Noaa2 a zdroje oblačnosti Energocentrum. Bohužel zdroj Energocentrum je placený, tedy pro finální algoritmus nepoužitelný. S uvážením faktu, že některé zdroje jsou z komerčního hlediska nedostupné



vychází jako nejlepší kombinace oblačnosti z Noaa1 a zdroj osvitu Noaa2, které při uvedení zdroje v případě Noaa1 jsou aplikovatelné pro komerční účely.

Specifickým případem je potom zdroj oblačnosti Noaa2, který je stahován teprve od začátku roku, a proto je potřeba jej vyhodnotit zvlášť. Pro posouzení výstupů oblačnosti Noaa2 je po odečtení množiny dat potřebné na učení k dispozici pouze 18 dní, tedy toto vyhodnocení se nedá pokládat za směrodatné. Data, pro která byl zdroj testován jsou v rozsahu od 12. 4. 2017 do 2. 5. 2017.

Tabulka 6: Srovnání zdrojů pro Noaa2

RMSE		zdroj osvitu			
		Hawai	Medard	Noaa1	Noaa2
zdroj oblačnosti	Eng	282,2	287,6	289,6	293,9
	IO	277,2	291,0	278,9	287,9
	Noaa1	277,5	290,5	278,1	288,1
	Noaa2	308,4	317,6	310,2	321,8

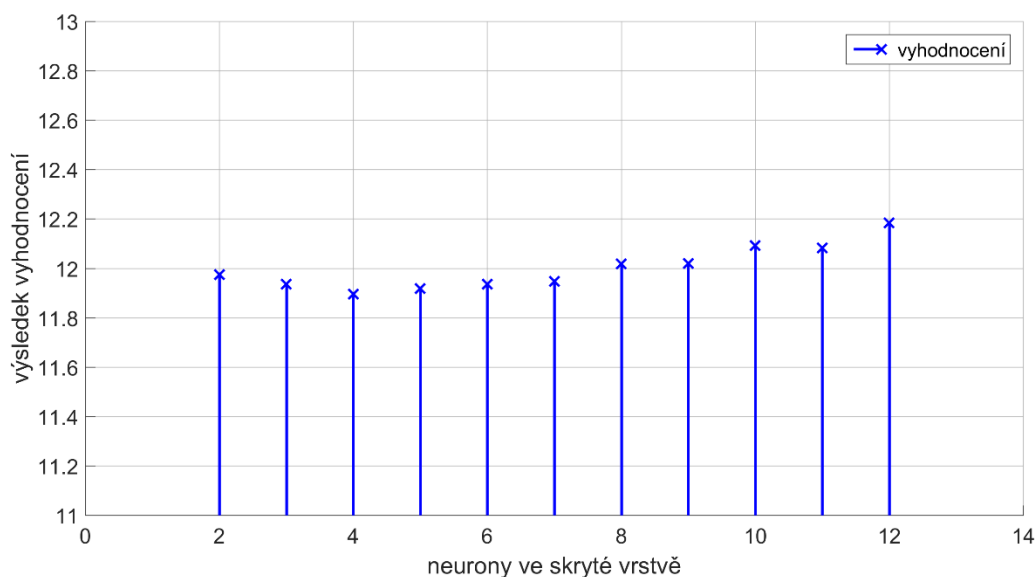
Při posouzení výsledků z tabulky 5 se zdroj oblačnosti Noaa2 ve srovnání s konkurencí nezdá příliš kvalitní, nicméně při analýze na větším počtu dat se situace může změnit. Zdroj také není posunutý tak, jako ostatní zdroje viz 7.2 *Úpravy posunutí dat vůči realitě*.

Zásadní změna situace kvality některých zdrojů je zřejmě způsobena vícero faktory. Hlavním z nich je, že vyhodnocení z podkapitoly 7.4 *Srovnání chyby sítě s různými zdroji osvitu* bylo počítáno s jiným preprocesingem oblačnosti a nepředpokládalo se, že to bude mít tak zásadní vliv.

Starší vyhodnocení zdrojů s původním preprocesingem je součástí článku *Přesnost zdrojů předpovědi osvitu pro fotovoltaické systémy* [28], kde jsou vypouštěny noční hodnoty, v závislosti na měřených datech a dominuje zpoplatněný zdroj SODA. V závěru práce bude srovnání provedeno stejným způsobem, a bude tak možné výstupy porovnat.

## 9 Predikce výkonu na FV elektrárně Unhošť

Neuronová síť navržená pro predikci na elektrárně Unhošť má stejný tvar jako pro osvit, viz obr. 23. Vzhledem ke změně předpovídaného výstupu by v ideálním případě bylo dobré udělat znovu odladění množství dat potřebného k učení a počtu neuronů ve skryté vrstvě. První úloha je časově velmi dlouhá, takže se pro tento problém počítá, že množství dat se shoduje. Z výstupu hledání optimálního počtu neuronů ve skryté vrstvě je patrné, že sítě budou velmi podobné a předpokládaná chyba způsobená tímto nedostatkem nebude velká. Úloha je spočítána od 17.5.2016 do 5.4.2017. Zdroje jsou zvoleny, na rozdíl od všech předešlých úloh podle výstupu z předchozí kapitoly (8.4 Závěrečné srovnání zdrojů).



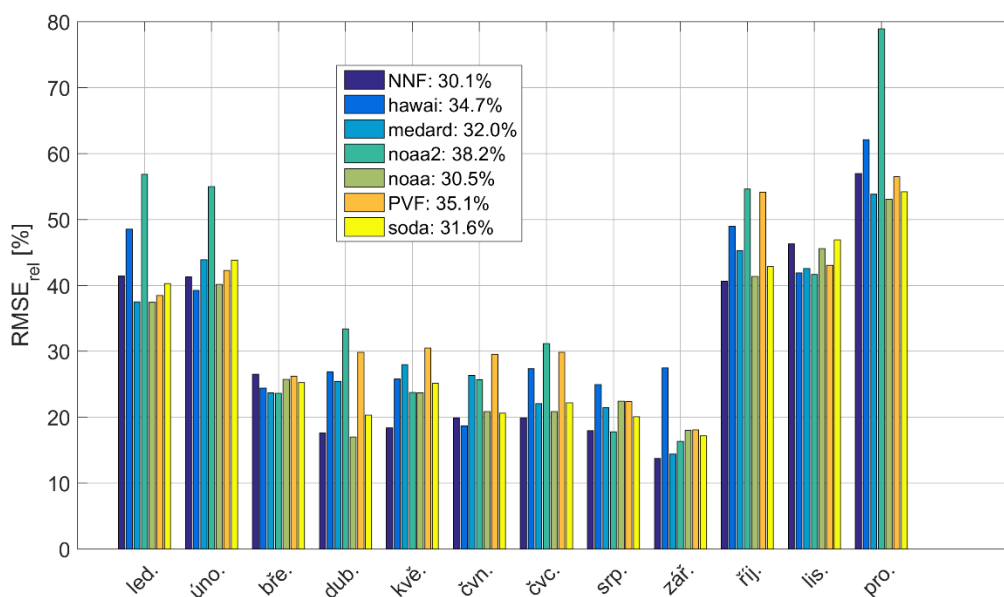
Obrázek 32: Hledání vhodného počtu neuronů pro předpověď výkonu Unhošť

Z grafu na obr. 32 je patrné, že nejmenší chyba bude na 4 neuronech, chování v závislosti na počtu neuronů je velmi podobné jako v případě předpovědi osvitu. Data jsou trénována na kratším časovém intervalu a měsíce květen a duben, které jsou počítány jen pro menší počet dní, mohou působit drobnou odchylku, protože obecně platí, že pro každý měsíc je vhodnější jiná síť.

## 10 Zpracované výstupy předpovědi osvitů

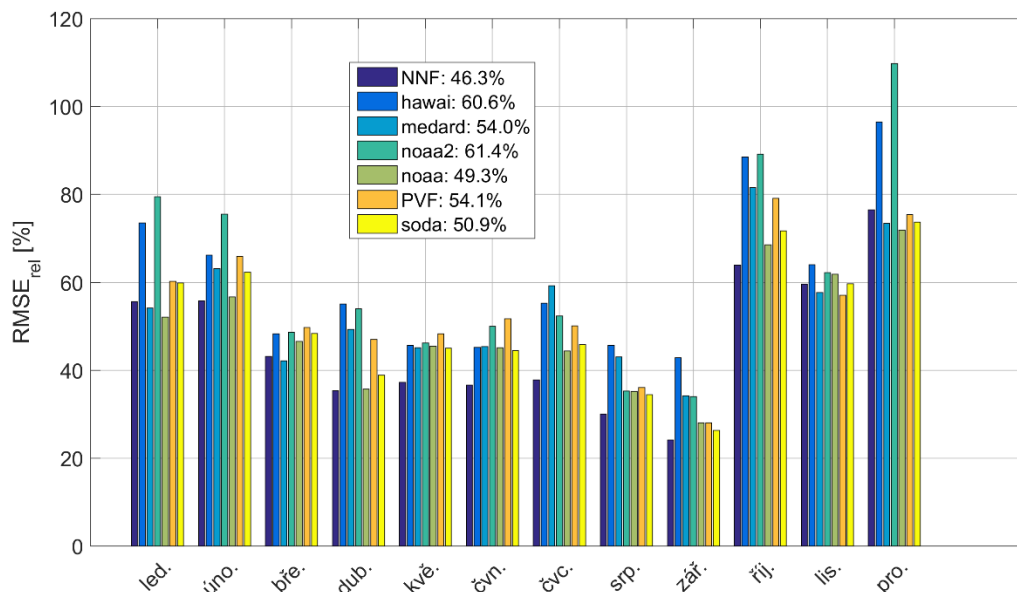
Následující grafy ukazují chyby výstupů algoritmu vůči naměřeným hodnotám ve srovnání s PV Forecast a dostupnými zdroji. Chyby v grafech této kapitoly jsou počítány jako chyby hodinových hodnot, nebo denních sum v jednotlivých měsících. Ve výpočtu nejsou zahrnuty noční hodnoty, což se uvažuje podle měřených dat.

Je možné, že zajímavější by bylo určovat je ne striktně podle naměřených dat, ale zvážit i hodnoty předpovědí a zahrnout do výpočtu i situaci, kdy předpověď nesedí s východem či západem slunce. Takto může nastat situace, kdy předpověď má o dvě denní hodnoty víc, ale na výpočtu chyby dat se to neprojeví. Nicméně tento způsob vyhodnocování je použit v práci [Benchmarking of Different Approaches to Forecast Solar Irradiance](#) [23] i v [Přesnost zdrojů předpovědi osvitů pro fotovoltaické systémy](#) [28] a pro srovnání má smysl se od něho neodklánět.

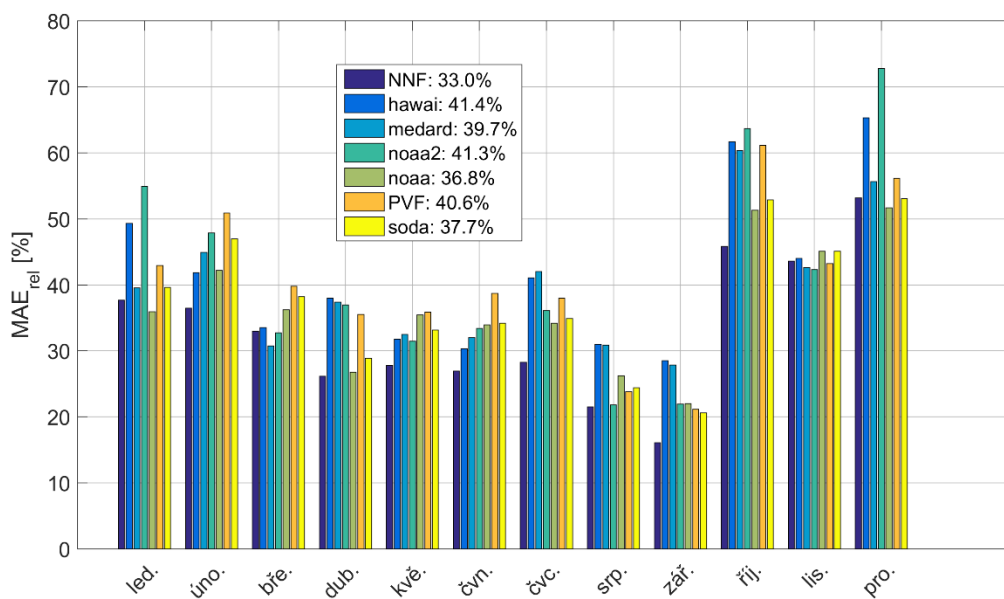


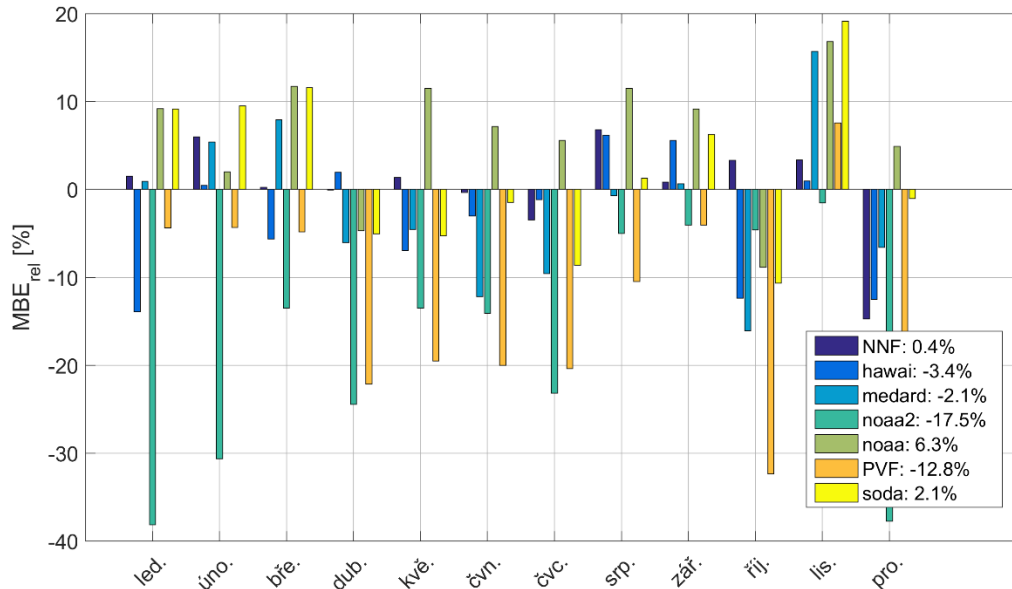
Obrázek 33: Srovnání predikcí,  $MRSE_{rel}$  pro denní sumy

Na grafech z obr. 33 a 34 je vidět, že zdroj Noaa1 (v grafu jako noaa) je v rámci konkurence poměrně dobrý, ačkoliv z předchozích analýz dat, které zahrnují noční hodnoty viz 8.4 Závěrečné srovnání zdrojů lze nabýt opačného dojmu. Z toho plyne, že je pravděpodobně možné na preprocessingu těchto dat ještě zapracovat.

Obrázek 34: Srovnání predikcí,  $MRSE_{rel}$  pro hodinové hodnoty

Srovnání v grafu na obr. 34 ukazuje, že ve studenějších měsících je obtížnější osvit předpovídat. Jedním z možných zlepšení je zkusit najít síť vhodnou pro zimní měsíce, protože, jak lze vypočítat právě v zimních měsících, je predikce sítě v rámci konkurence o něco horší, zatímco v letních dominuje. Zde je dobré připomenout, že to není způsobeno větším vlivem chyb MAE a RMSE v letních měsících, protože síť byla záměrně nastavována s přihlédnutím k této skutečnosti.

Obrázek 35: Srovnání predikcí,  $MAE_{rel}$  pro hodinové hodnoty

Obrázek 36: Srovnání predikcí,  $MBE_{rel}$  pro hodinové hodnoty

Z grafu na obr. 36 plyne tendence zdrojů k vychylování ve směru osy  $y$ . V tomto ohledu je na tom predikce neuronové sítě poměrně dobře.

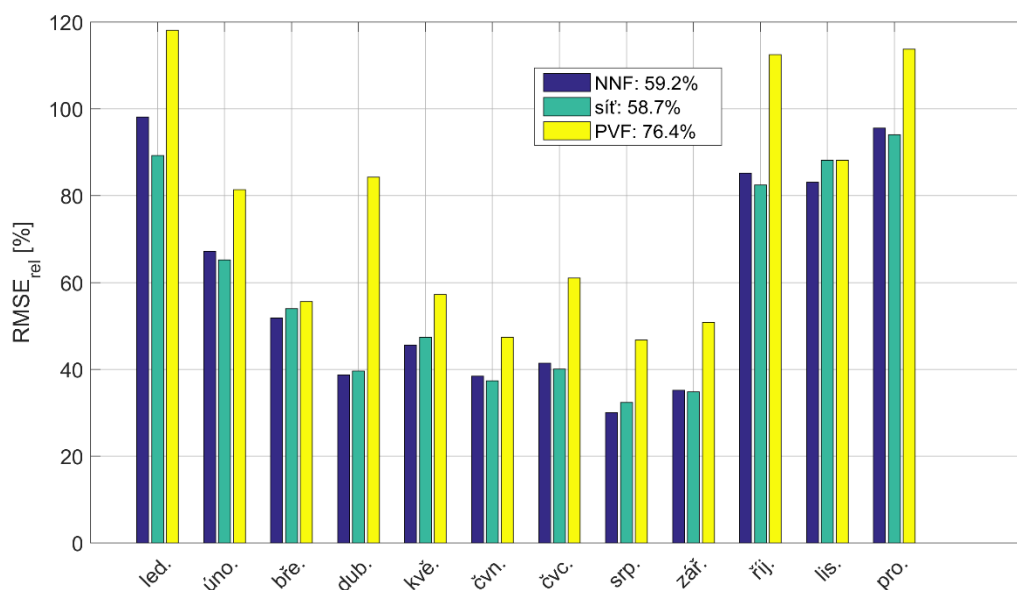
Grafy celkově vypovídají o tom, že nasazení neuronové sítě by mělo být přínosné zejména v letních měsících. Pro veškerá celoroční zhodnocení vychází neuronová síť vítězně, ačkoliv pro jednotlivé měsíce tomu tak vždy být nemusí.

Podle [23] je obecně mnohem obtížnější předpovídat pro interkontinentální klima, než například pro přímořské oblasti jako je Španělsko apod. U středoevropského klimatu je dobré, pokud se hodnoty  $RMSE_{rel}$  pohybují v rozsahu 40-60 %, kam se neuronová síť s hodnotou 46,3 % s přehledem vejde.  $RMSE_{rel}$  předpovědi pro Španělsko se naproti tomu pohybují mezi 20 a 35 %.

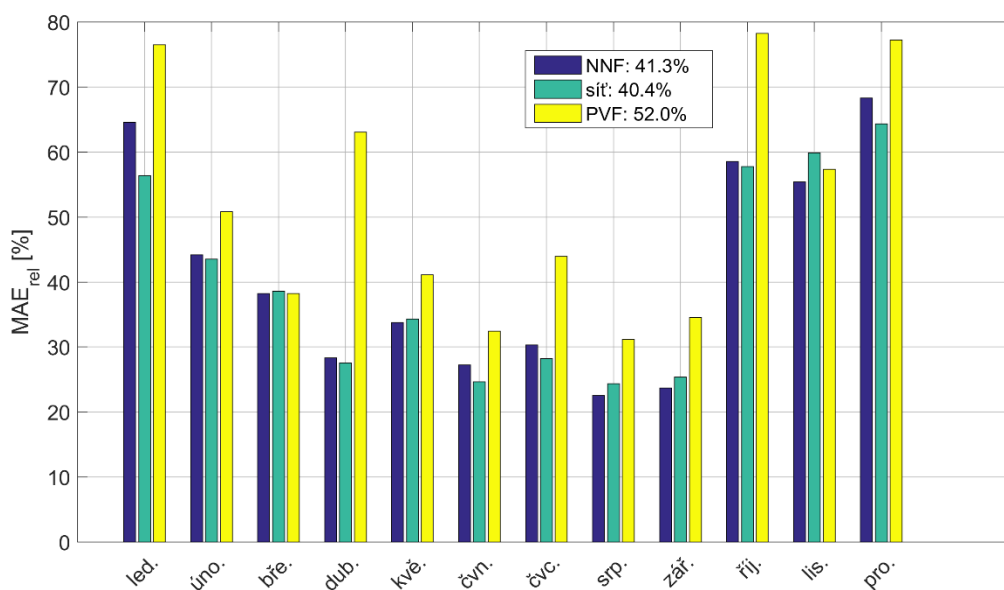
Grafy se zbývajícími chybami budou na příloženém médiu ve složce *final evaluation*.

## 11 Zpracované výstupy předpovědi výkonu

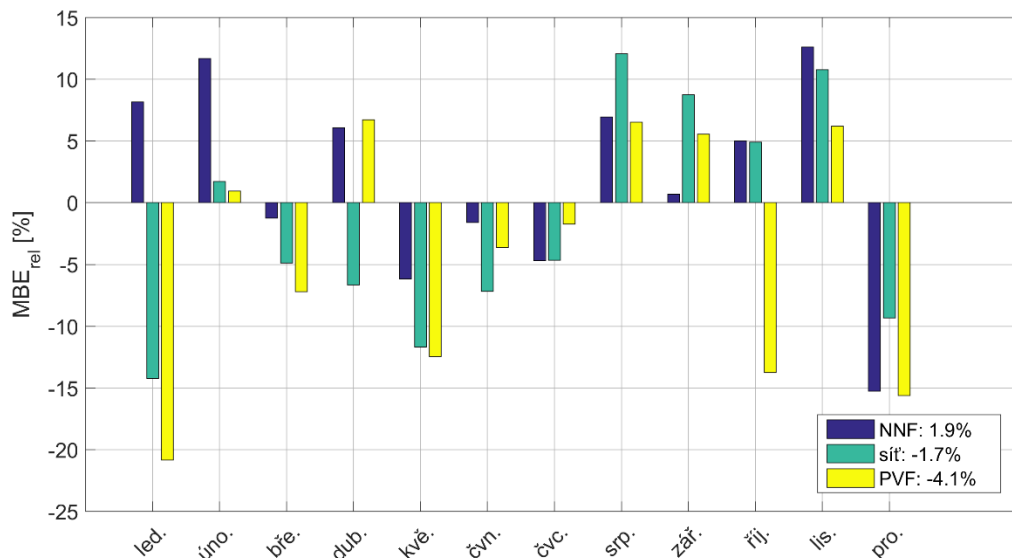
Kromě neuronové sítě pro předpověď osvětlení mělo být součástí práce také srovnání s konvenčním modelem. Ten předpokládá využití teploty v kombinaci s osvětlením. Bohužel se ukazuje, že současná metoda teplotu nevyužívá, a v kontextu služeb poskytovaných na ČVUT UCEEB by takové srovnání nedávalo smysl. Při znalosti poskytovaných služeb se jeví smysluplnější spíše porovnání se současným přepočtem na výkon, který je využit v budově Fénix v Jeseníku, než s modelem, který se v kombinaci se službou nikde nepoužívá.



Obrázek 37: Vyhodnocení predikce výkonu  $RMSE_{rel}$ , PVF, NNF a network odpovídají variantám 1,2 a 3 z textu pod grafy



Obrázek 38: Vyhodnocení predikce výkonu  $MAE_{rel}$ , PVF, NNF a network odpovídají variantám 1,2 a 3 z textu pod grafy



Obrázek 39: Vyhodnocení predikce výkonu  $MBE_{rel}$ , PVF, NNF a network odpovídají variantám 1,2 a 3 z textu pod grafy

Grafy na obr. 37,38 a 39 tedy ukazují srovnání chyby předpovědi pomocí neuronových sítí se v současnosti používaným modelem. Model využívá fitování polynomem čtvrtého řádu metodou nejmenších čtverců, kde vstupem je jen osvit. V grafech je srovnání následujících variant:

- 1) *data* → *Předpověď osvitu pomocí PV Forecast* → *predikovaný výkon v kW (PVF model)*
- 2) *data* → *Předpověď osvitu pomocí NN Forecast* → *predikovaný výkon v kW (NNF model)*
- 3) *data* → *Neuronová síť pro předpověď výkonu* → *predikovaný výkon v kW*

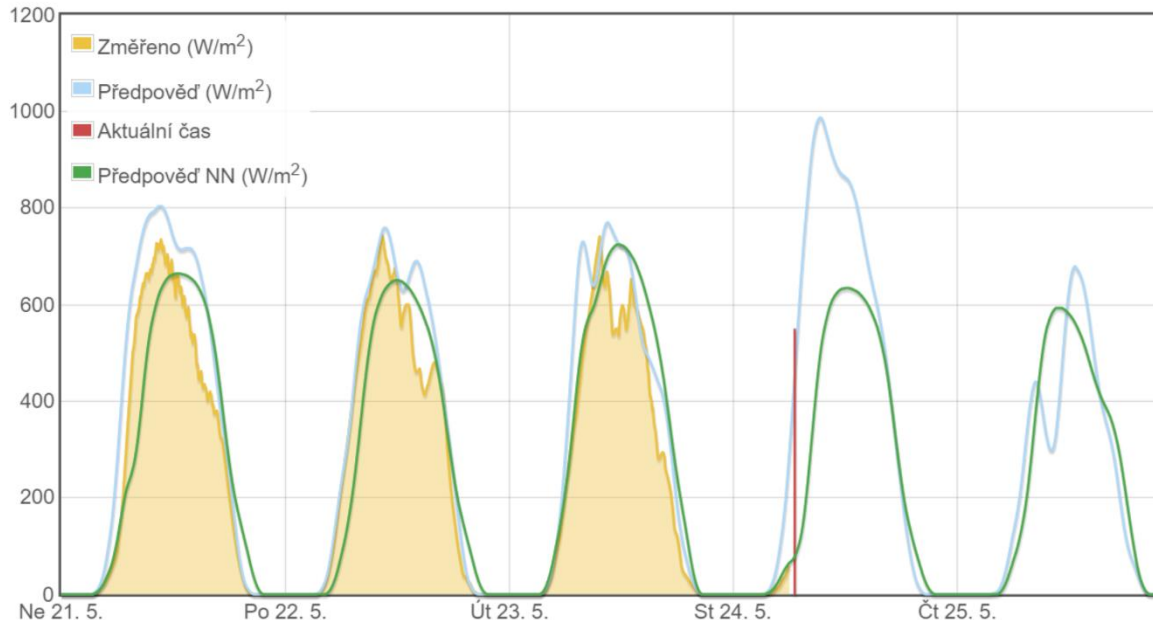
I v případě predikce výkonu lze ve spojitosti se sítí hovořit o úspěchu. Ačkoliv v tomto případě není náskok sítě tak zásadní. Je možné, že zde půjde o nějakou chybu měření či zpracování dat, protože v měsících s kratšími dny je chyba vysoká pro všechny varianty. Síť by teoreticky tímto problémem mohla být postižena více, protože další dva modely jsou trénovány pro každý měsíc zvlášť, kdežto neuronová síť obecně stále vyžaduje větší trénovací množinu dat. To, že je síť i tak o trochu lepší, ukazuje značný potenciál aplikace neuronové sítě i na tuto úlohu.

## 12 Závěr

Pro odladění sítě bylo potřeba nejprve upravit skripty pro stažení a zpracování dat. Bohužel se tak velkou součástí práce stal nový preprocessing dat, který zabral odpovídající část času, a ještě se do něj určitě vyplatí investovat. Ukázalo se, že přesnost jednotlivých zdrojů je velmi závislá právě na preprocessingu dat. S novými skripty bylo dosaženo diametrálně odlišných výsledků, než tomu bylo před příslušnými úpravami a obecně se dá říci, že kvalita nově upravených dat je vyšší (Viz kapitola 7 *Navržená řešení pro preprocessing dat*).

Ve srovnání s ostatními zdroji momentálně neuronová síť vyniká (viz 10 *Zpracované výstupy předpovědi osvitů*) a dá se očekávat, že s lepším zpracováním dat a zapracováním poznámek z kapitoly 13 *Budoucí práce* bude kvalita i nadále o krok před v práci zmíněnými predikcemi. Mimo potřebných dat a času, které v současné době nejsou k dispozici, bude také zapotřebí nemalý početní výkon.

V souladu se zadáním byla neuronová síť pro předpověď osvitů implementována na server a příslušná verze algoritmu je nahrána na přiložené médium (Část kódu je kvůli bezpečnosti odstraněna). Jak jsou výstupy sítě prezentovány na serveru je vidět na obr. 40.



Obrázek 40: Náhled výstupu sítě na serveru

Implementace sítě je v experimentálním režimu a zatím jen pro jeden bod viditelný pouze při internetovém připojení přímo na ČVUT UCEEB. V budoucnu by mělo dojít k rozšíření na plochu celé České republiky, ačkoliv k tomu bude pravděpodobně zapotřebí zvýšit výpočetní výkon serveru.



Kapitola 11 zpracované výstupy předpovědi výkonu ukazuje na další směr aplikace neuronových sítí pro forecasting. Síť z této kapitoly práce je také poměrně dobře odladěna, nicméně stejně jako v případě predikce osvitů je u ní ještě potenciál ke zlepšení.

Práce na službě PV Forecast touto prací rozhodně nekončí. Vzhledem k přístupu týmu, který se o službu stará a pokroku pozorovaném během období, kdy byla tato práce psána, není nejmenší pochybnost o dalším vývoji služby. V poslední kapitole této práce, bude nastíněno několik kroků, kterými lze navázat.

## 13 Budoucí práce

Vzhledem k nemalým možnostem neuronových sítí, neoptimálním preprocessingu některých zdrojů dat a dalším problémům se předpokládá další a patrně rozsáhlejší pokračování práce na algoritmu. V této kapitole budou navrženy některé způsoby, jak výsledný algoritmus rozšířit a zlepšit.

### 13.1 Zpracování a stahování dat

Obecně lze říci, že původní skripty pro stahování dat nejsou v dobrém stavu. Vzhledem k tomu, že to nebyl hlavní záměr této práce, nebylo v časových možnostech se do této problematiky ještě více ponořit.

#### 13.1.1 Osvit

Výstupy při posunování dat ukazují, že pro jednotlivé měsíce se stále vyplatí data posunout. Pochopitelně drobné odchylky mohou být následkem špatných předpovědí, ale obecně by tomu tak být nemělo. Pro posouzení takových dlouhodobých trendů je dobré použít data z delších časových období, než jaká jsou momentálně k dispozici.

Dalším krokem je provést lepší rozbor zdrojových dat. Zejména rozbor příčin posunutí časových značek apod. by mohl být klíčem ke zlepšení předpovědí a nalezení důvodů proč k posunutí dochází.

Minimálně pro zdroj Noaa1 pravděpodobně bude možné udělat lepší preprocessing.

#### 13.1.2 Oblačnost

Pro zdokonalení preprocessingu oblačnosti je zjevně ještě více prostoru než v případě osvitu. Nepochybně nejdůležitějším krokem je vyhodnocení oblačnosti ze zdroje Noaa2, jakmile bude k dispozici větší rozsah dat.

Stejně jako v případě osvitu i zde by bylo dobré k posouzení dlouhodobých trendů zpracovat data z delšího časového období.

Dále je vhodné provést početně mnohem náročnější a podrobnější analýzu oblačnosti v intervalu po 1 minutě, následně grafy proložit rozumnými křivkami a vyhodnotit posuny s přihlédnutím ke konkrétní lokaci (Viz konec podkapitoly 7.2 *Úpravy posunutí dat vůči realitě*).

Na rozdíl od osvitů, kde jsou naměřená data z některých lokací, pro oblačnost je potřeba taková data teprve získat. V lokaci ČVUT UCEED byl instalován nový sky imager<sup>vii</sup>, takže tato možnost po nasbírání většího množství dat nastane.

## 13.2 neuronová síť – predikce osvitů

Neuronová síť použitá v algoritmu bezpochyby nějakým způsobem funguje. Vzhledem k tomu, že proniknout do problematiky, upravit data a vyladit síť je časově náročná operace, je v práci uvažována jen jednovrstvá síť. Nabízí se tedy možnost experimentovat se sítí o 2 vrstvách, sítí s jinými aktivačními funkcemi, dalšími vstupy (např. teplotou) apod. Pochopitelně potenciál neuronových sítí je značný a možností, jak ji nastavit je celá řada.

Stejně jako výše, je zde možný přínos přepracování vyvážení sítě po získání delšího intervalu dat. Vyvažování sítě v kapitole 8.2 *Hledání vhodné topologie sítě* se uvažuje jako základní nastavení. V případě zpracování dat z několika let zpětně půjde dosáhnout toho, že tak jako se mezi jednotlivými měsíci mění počet vzorků pro učení sítě, bude se měnit i počet neuronů ve skryté vrstvě. Například v datech z roku 2016, je optimální počet neuronů pro leden 2, zatímco pro listopad více než 18.

## 13.3 neuronová síť – predikce výkonu

Pro síť použitou pro přepočítání Unhoště je potřeba udělat vícero kroků, než na kolik zbyl prostor. Data pro vyhodnocení předpovědi po odečtení učící množiny nezahrnují ani průběh celého roku. Síť je určena spíše k posouzení přínosu neuronové sítě pro přímou predikci výkonu bez nutného mezikroku, kterým je předpověď osvitů. Nicméně se dá předpokládat, že postupy a postřehy popsané v této práci, které se týkají predikce osvitů se dají použít i pro optimalizaci této sítě.

## 13.4 Nowcasting

Vzhledem k pokroku, jaké dosáhl NN Forecast oproti PV Forecastu je pravděpodobné, že aplikace neuronové sítě pro nowcasting by mohla být zajímavá. Výstupy výběru sítě metodou „ideal“ viz kapitola 8.1 *Trénování neuronové sítě* ukazují, že je mezi natrénovanými sítěmi je často jedna velmi dobrá. Pro nowcasting je po několika ranních hodinách k dispozici poměrně zajímavý vzorek dat a aplikací této informace by se zřejmě dalo hledat v průběhu dne vhodně natrénovanou síť.

---

<sup>vii</sup> Zařízení pro měření zakrytí oblohy.

V této práci nebylo využití sítí pro nowcasting uvažováno zejména kvůli nevhodně zpracovaným datům a náročnosti v zadání definovaných úloh. Také je nejdříve potřeba přesně určit lokální posuny v čase tak, jako to bylo nastíněno v kapitole 7.2 *Úpravy posunutí dat vůči realitě*, protože se předpokládá, že data nebudou vzorkována po hodinách, ale v řádu minut.

## Vysvětlivky





PV – fotovoltaický


NN – neuronová síť

forecast – dlouhodobá předpověď (v řádu hodin – dní)

nowcast – krátkodobá předpověď (v řádu minut – hodin)

## Seznam použitých zdrojů

- [1] Zpřesnění predikce výkonu fotovoltaické elektrárny (Projekt I),  
Zbyněk Bambušek  
*ČVUT – FEL, Technická 2, 166 27 Praha 6 – Dejvice* 
- [2] Zpřesnění predikce solárního ozáření (Projekt II),  
Zbyněk Bambušek  
*ČVUT – FEL, Technická 2, 166 27 Praha 6 – Dejvice* 
- [3] <http://oenergetice.cz> [online]  
Instalovaný výkon solárních elektráren v ČR od roku 2012 stagnuje, [info@oenergetice.cz](mailto:info@oenergetice.cz),  
využívá zpravodajství ČTK
- [4] Roční zpráva o provozu ES ČR 2015  
Oddělení statistiky a sledování kvality ERÚ, Praha 2016 
- [5] [https://is.mendelu.cz/eknihovna/opory/zobraz\\_cast.pl?cast=21471](https://is.mendelu.cz/eknihovna/opory/zobraz_cast.pl?cast=21471) [online]  
Mendelova univerzita v Brně, Zemědělská 1/1665, 613 00 Brno
- [6] Predikce krátkovlnného slunečního záření pro malé fotovoltaické systémy  
Martin Procházka  
*ČVUT – FEL, Technická 2, 166 27 Praha 6 - Dejvice*
- [7] AZE-P1-slunečni energie, doc. Ing. Tomáš Matuška, Ph.D.  
*ČVUT – fakulta strojní, Technická 2, 166 27 Praha 6 – Dejvice* 
- [8] About Forecast IO [Online]:  
<https://darksky.net/forecast/50.0833,14.4667/si24/en>
- [9] About Json format [Online]:  
<http://www.json.org/json-cz.html>
- [10] *About grib file format* [Online]:  
<http://www.wmo.int/pages/prog/www/DPS/FM92-GRIB2-11-2003.pdf>  
[Přístup získán 21. 5. 2016].
- [11] *About Noaa* [Online]: <http://www.noaa.gov/about-our-agency>  
[Přístup získán 4. 5. 2017].
- [12] *Global Climate and Weather Modeling Branch: Mission Statement* [Online]:  
<http://www.emc.ncep.noaa.gov/gmb/mission.html>  
[Přístup získán 4. 5. 2017].
- [13] *Global Forecast System (GFS) documentation* [Online]:  
[http://nomads.ncep.noaa.gov/txt\\_descriptions/GFS\\_doc.shtml](http://nomads.ncep.noaa.gov/txt_descriptions/GFS_doc.shtml) [Přístup získán 4. 5. 2017].

- [14] *Informace o modelu „Noaa 1“* [Online]:  
[http://nomads.ncep.noaa.gov:9090/dods/fnl/fnl20160409/fnlflx\\_00z.info](http://nomads.ncep.noaa.gov:9090/dods/fnl/fnl20160409/fnlflx_00z.info)  
 [Přístup získán 4. 5. 2017]. Poznámka: do odkazu je nutné vložit aktuální datum
- [15] *Informace o modelu „Noaa 2“* [Online]:  
[http://nomads.ncep.noaa.gov:9090/dods/gfs\\_0p25/gfs20160331/gfs\\_0p25\\_00z.info](http://nomads.ncep.noaa.gov:9090/dods/gfs_0p25/gfs20160331/gfs_0p25_00z.info) [Přístup získán 4. 5. 2017]. Poznámka: do odkazu je nutné vložit aktuální datum
- [16] *Informace o zdroji Soda* [Online]:  
[http://www.soda-is.com/eng/about/soda\\_project\\_factsheet.html](http://www.soda-is.com/eng/about/soda_project_factsheet.html) [Přístup získán 4. 5. 2017].
- [17] *Služby poskytované zdrojem Soda* [Online]:  
<http://www.soda-is.com/eng/services/index.html> [Přístup získán 4. 5. 2017].
- [18] *Profil společnosti ENERGOCENTRUM PLUS, s.r.o.* [Online]:  
<http://www.energocentrum.cz/profil-spolecnosti/> [Přístup získán 4. 5. 2017].
- [19] *Předpověď počasí společnosti ENERGOCENTRUM PLUS, s.r.o.* [Online]:  
<http://rcware.eu/weather/prague/> [Přístup získán 4. 5. 2017].
- [20] *Návoděda služby Medard* [Online]: <http://www.medard-online.cz/help>  
 [Přístup získán 4. 5. 2017].
- [21] *About PacIOOS* [Online]: <http://oos.soest.hawaii.edu/pacioos/about/>  
 [Přístup získán 4. 5. 2017].
- [22] *Data Access Form - Hawaii* [Online]:  
[http://oos.soest.hawaii.edu/erddap/griddap/NCEP\\_Global\\_Best.html](http://oos.soest.hawaii.edu/erddap/griddap/NCEP_Global_Best.html)  
 [Přístup získán 4. 5. 2017].
- [23] Benchmarking Of Different Approaches To Forecast Solar Irradiance  
 Elke Lorenz<sup>1</sup>, Jan Remund<sup>2</sup>, Stefan C. Müller<sup>2</sup>, Wolfgang Traunmüller<sup>3</sup>, Gerald Steinmaurer<sup>4</sup>,  
 David Pozo<sup>5</sup>, José Antonio Ruiz-Arias<sup>5</sup>, Vicente Lara Fanego<sup>5</sup>, Lourdes Ramirez<sup>6</sup>, Martin  
 Gaston Romeo<sup>6</sup>, Christian Kurz<sup>7</sup>, Luis Martin Pomares<sup>8</sup>, Carlos Geijo Guerrero<sup>9</sup> University of  
 Oldenburg, Institute of Physics, Energy and Semiconductor Research Laboratory, Energy  
 Meteorology Unit, Carl von Ossietzky Strasse 9-11, 26129 Oldenburg, Germany,  
 elke.lorenz@uni-oldenburg.de
- 1) Meteotest, Fabrikstrasse 14, 3012 Bern, Switzerland
  - 2) Blue Sky Wetteranalysen, Steinhüblstr.1, 4800 Attang Puchheim, Austria
  - 3) ASiC – Austria Solar Innovation Center, Roseggerstraße 12, 4600 Wels, Austria,
  - 4) University of Jaén, Department of Physics, Campus Lagunillas, 23071, Jaén, Spain
  - 5) CENER, Ciudad de la Innovación 7, 31621 Sarriguren (Navarre), Spain
  - 6) Meteocontrol GmbH, Spicherer Straße 48, 86157 Augsburg, Germany
  - 7) Ciemat, Energy Department, Av. Complutense 22, 28040, Madrid, Spain
  - 8) AEMet, Leonardo Prieto Castro, 28071, Madrid, Spain
- 
- [24] *Korelační a regresní analýza* [online]:  
[https://wikisofia.cz/wiki/Korela%C4%8Dn%C3%AD\\_a\\_regresn%C3%AD\\_anal%C3%BDza](https://wikisofia.cz/wiki/Korela%C4%8Dn%C3%AD_a_regresn%C3%AD_anal%C3%BDza)  
 [Přístup získán 4. 5. 2017].

[25] *Dokumentace k programu MATLAB* [online]:  
<https://www.mathworks.com/help/matlab/>, případně zavolání metody *help.m* přímo v programu

[26] Solar irradiation forecasting: state-of-the-art and proposition for future developments for small-scale insular grids

Hadja Maïmouna Diagne, *Réuniwatt, 14, rue de la Guadeloupe 97490 Sainte-Clotilde France*  
Mathieu David, Philippe Lauret, *University of La Reunion, 97715 Saint Denis Cedex 9*  
John Boland, *University of South Australia*



[27] Využití předpovědi osvitů pro řízení hybridních fotovoltaických systémů

J. Včelák, P. Wolf, J. Šulc, M. Procházka, M. Maška, J. Kohout  
ČVUT – Univerzitní centrum energeticky efektivních budov, *Třinecká 1024, Buštěhrad*

[28] Přesnost zdrojů předpovědi osvitů pro fotovoltaické systémy (článek ještě nevyšel, měl by vyjít v časopise Vytápění větrání instalace v létě 2017)

Ing. Petr Wolf, Ph.D., Ing. Jan Včelák, Ph.D., Ing. Tomáš Bäumelt, Erik Novák, MSc.,  
Ing. Jan Kohout  
ČVUT – Univerzitní centrum energeticky efektivních budov, *Třinecká 1024, Buštěhrad*



[29] *Služba PV Forecast* [online]:

<http://www.pvforecast.cz/predpoved-body.php>. [Přístup získán 17. 5. 2017].  
ČVUT – Univerzitní centrum energeticky efektivních budov, *Třinecká 1024, Buštěhrad*



## Obsah přiloženého média

- Tato práce
- Všechny zdroje ze seznamu s ikonou diskety
- Skripty na výpočet a vyhodnocení různých bodů práce
- Skripty na preprocessing a stažení dat
- Finální verze algoritmu (serverová a lokální)