



**Používateľská Príručka Systému**  
**pre Estimáciu Produkcie Fotovoltickej Energie**

Peter Krammer, Marcel Kvassay, Ondrej Habala, Ján Mojžiš, Ladislav Hluchý  
{peter.krammer, marcel.kvassay, ondrej.habala, jan.mojzis, ladislav.hluchy}@savba.sk

Verzia príručky: 1.0c  
Október 2023, Bratislava

<b>Obsah</b>	
<b>1.1 Požiadavky systému</b>	<b>2</b>
<b>1.2 Konfigurácia systému</b>	<b>3</b>
<b>1.3 Použitie systému</b>	<b>5</b>
<b>1.3.1 Použitie konzolového prístupu</b>	<b>7</b>
<b>1.3.2 Použitie grafického rozhrania</b>	<b>9</b>
<b>1.3.3 Používanie modulov systému z programovacieho jazyka Java</b>	<b>11</b>
<b>1.4 Štruktúra súborov</b>	<b>13</b>
<b>1.4.1 Konfiguračný súbor settings.ini</b>	<b>13</b>
<b>1.4.2 Vstupné dátové súbory</b>	<b>14</b>
<b>1.4.3 Súbory Číselníkov</b>	<b>15</b>
<b>1.4.4 Výstupné súbory</b>	<b>17</b>
<b>1.4.5 Previazanie dát</b>	<b>18</b>
<b>1.5 Prehľad chybových kódov a spôsoby odstránenia chýb</b>	<b>19</b>

## **1.1 Požiadavky systému**

Minimálne hardwarové požiadavky na prevádzku systému na tréning a predikovanie (estimáciu) úrovne generovanej fotovoltaickej energie (FVE) sú nasledovné:

- Min. 32 GB pamäte RAM
- Min. 4 GB voľného priestoru na pevnom disku

Doporučené hardwarové požiadavky na prevádzku systému:

- 64 GB pamäte RAM
- 8 GB voľného priestoru na pevnom disku
- Použitie rýchleho SSD disku pre adresár na ukladanie dočasných a výstupných súborov.

Uvedené hardwarové požiadavky sú testované pre 18 lokalít, pri celkovej počte cca 70000 záznamov určených na tréning / predikciu. Pri vyššej počte lokalít, resp. záznamov sa jednotlivé požiadavky môžu navyšovať (od minimálnych požiadaviek bližšie k doporučeným), primárne v závislosti od počtu lokalít.

Softwarové požiadavky: potreba nainštalovanej Java Virtual Machine s verziou zodpovedajúcej Java 1.8, alebo vyššej.

## 1.2 Konfigurácia systému

Konfigurácia systému vyžaduje:

- Inštaláciu Softwaru Java Virtual Machine (pre Javu 1.8 resp. vyššiu verziu) pre daný operačný systém (napr. z: <https://www.java.com/en/download/manual.jsp> )
- Nastavenie adresárov v súbore settings.ini (Štruktúra súboru je uvedená aj v kapitole 1.4)
- Vyplnenie číselníkov - ciselnik\_instal\_vykonu.csv a lokality\_FVE\_ciselnik.csv (v adresári číselníkov určenom v settings.ini) v súlade so špecifikáciou určenou v kapitole 1.4
- Nastavenie práv na umožnenie zapisovania do adresára dočasných súborov a do adresára výstupov

### Nastavenie súboru settings.ini

Tento súbor obsahuje 3 riadky s úplnými cestami k špecifickým adresárom:

- Prvý riadok určuje plnú cestu k adresáru, kde môže estimačný systém vytvárať dočasné pracovné súbory (čiže daný používateľ systému musí mať práva na zápis do tohto adresára; zároveň daný adresár musí byť umiestnený, na zariadení, kde je možné realizovať priamo zápis (teda nie DVD a pod.) a taktiež kde je dostatočný voľný diskový priestor).
- Druhý riadok určuje plnú cestu k adresáru obsahujúcemu potrebné súbory číselníkov. Patria sem súbory číselníkov: ciselnik\_instal\_vykonu.csv a lokality\_FVE\_ciselnik.csv Ich štruktúra je popísaná v časti 1.4.
- Tretí riadok určuje plnú cestu k adresáru, určenému pre výstupy. Do tohto adresára doporučujeme umiestňovať súbory natrénovaných modelov ako aj výstupy estimácie.

Na oddelovanie adresárov doporučujeme použiť bežné lomítka ( znak '/' ), ktorý funguje ako v operačnom systéme Linux, tak aj Windows. Taktiež je potrebné rozlišovať veľké a malé písmená v názvoch adresárov.

Príklad vyplnenia súboru settings.ini:

```
c:/Users/Peter/Documents/NetBeansProjects/FVE/tmpDir/  
c:/Users/Peter/Documents/NetBeansProjects/FVE/ciselniky/  
c:/Users/Peter/Documents/NetBeansProjects/FVE/vystupy/
```

*Odporúčanie: Pre zjednodušenie procesu odporúčame na lokálnom disku s dostatkom voľného miesta (min. 4 GB), v adresári, kde má používateľ plné práva, vytvoriť adresár FVE. V rámci neho vytvoriť nasledovné podadresáre:*

FVE/tmpDir/	- pre ukladanie dočasných súborov
FVE/ciselniky/	- pre ukladanie súborov číselníkov

FVE/vystupy/

- pre ukladanie výstupov (modelov, či estimovaných hodnôt)

FVE/data/

- pre umiestnenie dát v tvare csv súborov (FVE a meteo dáta)

FVE/dist/

- obsahujúci samotný estimačný systém vrátane knižníc

*Následne do adresára dist skopírovať samotný systém pre estimáciu FVE (jeho programové súbory vrátane jar súborov). V súbore settings.ini nastaviť úplné cesty na prvé tri vyššie uvedené adresáre.*

### 1.3 Použitie systému

Používanie systému je možné výhradne po realizovaní konfigurácie, v ktorej sú zabezpečené činnosti ako napr. nainštalovanie Java Virtual Machine, nastavenia adresárov a pod. Táto činnosť je popísaná v kapitole 1.2 – Konfigurácia systému.

Použitie takto nakonfigurovaného systému pozostáva z 2 fáz:

- A.) Fáza tvorby modelu (trénovania modelu)
- B.) Fáza estimácie hodnôt FVE

Vo fáze A. tvorby modelu je potrebné poskytnutie čo najrozsiahlejšieho objemu dát (pre maximálnu mieru reprezentatívnosti), ako vstupných dát (dátumy, časy, meteo-dáta) tak aj výstupných dát (príslušných hodnôt meranej FVE za dané obdobie) vo formáte csv súborov. Taktiež je potrebné zabezpečiť korektné vyplnenie číselníkov definujúcich inštalovaný výkon a FVE lokality. Cieľom fázy modelovania je vytvorenie regresného modelu, ktorý bude pokrývať lokality zastúpené v dodaných údajoch. Výstupom tejto fázy je súbor s príponou .model, ktorý reprezentuje znalosti, resp. závislosti v dátach – teda vzťahy medzi vstupnými údajmi a výstupnou cieľovou veličinou. Takto vytvorený model, pre 18 lokalít zaberá obvykle cca 600 MB. Pre zabezpečenie korektnej dátovej integrity, model obsahuje aj pôvodné číselníky použité pri trénovaní. Proces trénovania modelu je potrebné realizovať aspoň raz, pričom výsledný natrénovaný model odporúčame zálohovať (napr. napálením na DVD / zálohovať na USB kľúč).

Fáza B – estimácia hodnôt FVE, už vyžaduje ako vstup aj natrénovaný model.

Cieľom tejto fázy je realizovať odhad (estimáciu) výkonov FVE na základe súvisiacich veličín ako je dátum, čas, meteo-veličiny a podobne. Príslušné odhadnuté hodnoty budú zapísané do výstupného súboru vo formáte csv. Presná štruktúra tohto výstupného súboru je uvedená v kapitole 1.4.

Očakáva sa, že proces estimácie bude realizovaný výrazne častejšie než proces trénovania, nakoľko dobre natrénovaný model, ktorý dáva uspokojivú presnosť estimácií, tak môže byť úspešne používaný aj niekoľko rokov. V prípade kvalitne natrénovaného modelu, už preto nie je potrebné dlhšiu dobu realizovať Fázu A, keďže model dáva dobré výsledky. V takomto prípade sa môže napr. už iba pravidelne realizovať fáza B – estimácie hodnôt FVE pre zadané lokality, dátumy, časy a meteo-veličiny.

V prípade ak estimácia nedáva uspokojivé výsledky (sú napr. nepresné), tak je vhodné zopakovať proces trénovania modelu s určitými zmenami. Napr. s vyšším rozsahom vstupných dát (napr. za dlhšie obdobie), použiť očistené dáta, alebo zoskupením niektorých lokalít do spoločnej. Taktiež je možné nanovo presnejšie odhadnúť inštalovaný výkon. A s takýmito zmenami zopakovať proces trénovania.

Predložený estimačný systém na určenie objemu vyprodukovanej FotoVoltickej Energie (FVE) je možné používať 3 rôznymi spôsobmi:

- Ako konzolovú aplikáciu (popísané v časti 1.3.1 – použitie konzolového prístupu), ktorej parametre špecifikujú vstupné a výstupné súbory
- Grafické rozhranie (popísané v časti 1.3.2 – použitie grafického rozhrania), v ktorom bude používateľovi umožnené, zvoliť cez dialógové okná, ktoré súbory budú použité pre tréning, či estimáciu.
- Programované riešenie z jazyka Java (popísaného v časti 1.3.3 - Používanie modulov systému z programovacieho jazyka Java) pripojením príslušných jar (Java Archive) knižníc.

Všetky 3 uvedené spôsoby umožňujú tréning modelu, ako aj vykonanie estimácie použitím už-natréňovaného modelu.

### 1.3.1 Použitie konzolového prístupu

Tento spôsob použitia systému je vhodný najmä pre použitie programu v súčinnosti s inými softvérmi, avšak je vhodný aj pre používateľov (ktorý preferujú prácu v konzolovom prostredí). Teda nevyžaduje sa grafické prostredie (či už v systéme Windows, či Linux), postačuje konzolové okno s príslušnými právami.

**Pre realizáciu tréovania modelu**, použijete príkaz v tvare nasledovného príkladu (v adresári, kde sa nachádza súbor FVE.jar a tiež settings.ini):

```
java -Xmx19000M -cp FVE.jar fve.FVE_train  
"c:/Users/Peter/Documents/NetBeansProjects/FVE/data/FVE_2021+2022.csv"  
"c:/Users/Peter/Documents/NetBeansProjects/FVE/data/meteo_2021+2022.csv"  
"RF_model2.model"
```

Významy jednotlivých parametrov:

`-Xmx19000M` – parameter zabezpečujúci nastavenie pridelenej pamäte pre Java Virtual Machine; Doporučujeme hodnotu aspoň 19 GB (19000M) prípadne vyššiu, pre vysoký objem použitých dát.

`-cp FVE.jar fve.FVE_train` – technický parameter určujúci meno knižnice, meno triedy v projekte a meno metódy ktorá sa má vykonať (nemeniť!).

Posledné 3 parametre určujú úplnú cestu k súborom FVE dát, meteo dát (oba typu csv) a výstupný súbor modelu s príponou .model .

Štruktúra oboch csv súborov musí zodpovedať štruktúre uvedenej v časti 1.4 – Štruktúra súborov. Jednotlivé parametre sú oddelované medzerou.

**Pre vykonanie procesu estimácie** s už-natrénovaným modelom použijete príkaz podľa uvedeného príkladu (v adresári kde sa nachádzajú súbory FVE.jar a settings.ini):

```
java -Xmx19000M -cp FVE.jar fve.FVE_predict  
"c:/Users/Peter/Documents/NetBeansProjects/FVE/data/FVE_2021+2022.csv"  
"c:/Users/Peter/Documents/NetBeansProjects/FVE/data/meteo_2021+2022.csv"  
"c:/Users/Peter/Documents/NetBeansProjects/FVE/vystupy/RF_model2.model"  
"c:/Users/Peter/Documents/NetBeansProjects/FVE/vystupy/predikovane.csv"
```

Významy jednotlivých parametrov:

`-Xmx19000M` – parameter zabezpečujúci nastavenie pridelenej pamäte pre Java Virtual Machine; Doporučujeme hodnotu aspoň 19 GB (19000M) prípadne vyššiu, pre vysoký objem použitých dát.

`-cp FVE.jar fve.FVE_predict` – technický parameter určujúci meno knižnice, meno triedy v projekte a meno metódy ktorá sa má vykonať (nemeniť!).



Posledné 4 parametre určujú v poradí: úplnú cestu k súborom FVE dát, meteo dát (oba typu csv), súbor modelu, ktorý má byť na estimáciu použitý a výstupný súbor, do ktorého budú zapísané výsledné estimované hodnoty. Jednotlivé parametre sú oddelované medzerou. Štruktúry jednotlivých súborov sú popísané v časti 1.4.

### 1.3.2 Použitie grafického rozhrania

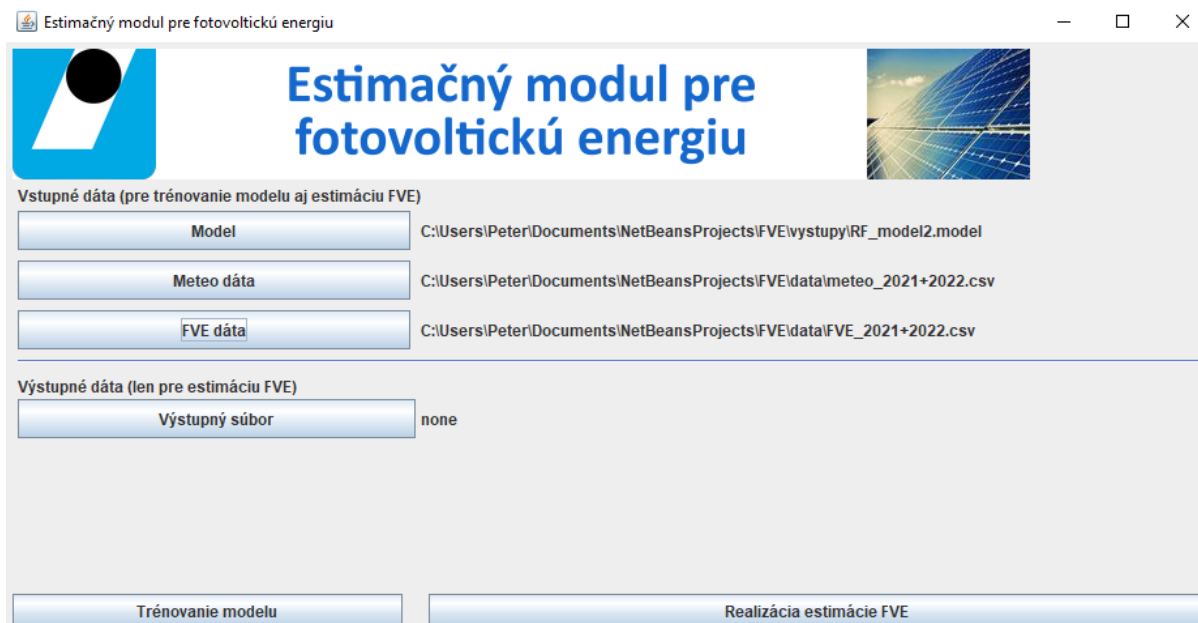
Tento spôsob použitia je vhodný pre bežných používateľov preferujúcich nakoľko nie je nutné vypisovať presné cesty a názvy súborov, ale je možné realizovať výber súboru dialógovým oknom – prechádzaním súborov na disku. Spustenie programu s grafickým rozhraním sa vykoná spustením súboru (v systéme Windows):

`spusti_gui.bat` - spustenie grafickej aplikácie  
resp. v systéme linux:  
`chmod 755 spusti_gui_linux.bat` - pre nastavenie práv na vykonanie  
`./spusti_gui_linux.bat` - pre spustenie grafickej aplikácie

#### Proces tréovania modelu

V procese tréovania modelu, musí používateľ zvoliť umiestnenie csv súborov s meteo-dátami a FVE dátami (zodpovedajúcim špecifikácií podľa časti 1.4). To sa deje po kliknutí na príslušné tlačidlo. Užívateľ taktiež zvolí umiestnenie a názov nového súboru modelu (treba mu dať príponu .model), ktorý bude v tomto procese vytvorený. Po vyplnení týchto 3 položiek sú zadané hodnoty viditeľné v pravej časti okna, vedľa jednotlivých tlačidiel.

Tréovanie modelu bude odštartované po stlačení tlačidla „Tréovanie modelu“ v ľavej dolnej časti obrazovky. Grafické znázornenie grafického rozhrania, na ktorom sú vidieť spomínané tlačidlá aj príklad vyplnených hodnôt s cestami a názvami súborov je znázornené na nasledovnom obrázku Obr. 1.



Obr. 1: Grafické znázornenie užívateľského rozhrania estimačného systému FVE; Na obrázku sú viditeľné tlačidlá pre zadávanie jednotlivých parametrov (v ľavej časti), ako aj aktuálne zvolené hodnoty parametrov (v pravej časti). Tlačidlá pre realizáciu tréovania, či estimácie sú v dolnej časti grafického okna.

Po stlačení tlačidla „Trénovanie modelu“ sa odštartuje proces prípravy dát a tréovania modelu, pričom tento proces bude prebiehať v samostatnom okne. Toto okno neuzatvárajte dokým výpočet nie je dokončený. Trénovanie modelu (pri počte lokalít 18) môže trvať cca 5-12 minút, v závislosti od použitia klasického, či SSD disku, rýchlosti CPU, ostatných bežiacich procesov a viacerých ďalších faktorov.

### **Proces estimácie výkonov FVE**

Analogický postup je aj v prípade realizácie estimácie. V tomto prípade je potrebné vyplniť všetky 4 parametre zodpovedajúcim 4 tlačidlám na ľavej strane grafickej obrazovky na obrázku Obr. 1 (teda aj vrátane 4. parametra - výstupného súboru, do ktorého budú zapísané výsledné estimované hodnoty FVE). Prvý uvedený parameter – model v tomto prípade slúži ako vstupný, teda musí sa odkazovať na existujúci súbor už-natrénovaného modelu.

Proces estimácie sa odštartuje tlačidlom v pravej dolnej časti grafického okna, ktoré je taktiež znázornené na obrázku Obr. 1. Opätovne je proces estimácie realizovaný v samostatnom okne, ktoré nesmie byť predčasne uzavreté.

V prípade ak vstupný FVE dátový súbor obsahoval aj skutočné hodnoty výkonov (ktoré nie sú vždy známe), tak na obrazovku budú vypísané informácie o dosiahnutej presnosti pre dané údaje. Dosiahnutá presnosť je vyjadrená použitím viacerých číselných charakteristík, ako napríklad: Pearsonov Korelačný koeficient, Spearmanov korelačný koeficient, Stredná absolútna chyba, Relatívna-Absolútna chyba a ďalších.

Vypočítané hodnoty estimovaného FVE výkonu sú uložené do súboru na disk, špecifikovaného ako posledný 4. parameter - výstupný súbor. Štruktúra tohto súboru je definovaná v kapitole 1.4.

### 1.3.3 Používanie modulov systému z programovacieho jazyka Java

Okrem priameho použitia predloženého estimačného systému, je možné daný systém začleniť aj do rozličných programových riešení, vyvíjaných v jazyku Java (od verzie 1.8) a to ako pre trénovací tak aj predikčný (estimačný) modul.

Pre úspešné začlenenie do iného programu, vyvíjaného v jazyku Java, je potrebné realizovať nasledovné kroky:

- pripojenie programových knižníc (FVE.jar, /lib/commons-math3-3.6.1.jar, /lib/jafama-2.3.2.jar, /lib/weka\*.jar ), do aktuálneho projektu. (napr. v prípade prostredia NetBeans, sa tento krok realizuje nasledovne: config -> customize -> Libraries -> Add Jar/Folder )

- Importovanie potrebných tried:

```
import fve.*;
```

- nakopírovanie konfiguračného súboru settings.ini do adresára v ktorom bude aktuálne spúšťaný projekt.

- Výrazne tiež doporučujeme nastavenie parametra javy pre zvýšenie objemu priradenej pamäte, pri jej spúšťaní príkazom -Xmx, napr. na hodnotu -19000M (resp. vyššiu hodnotu).

```
java -Xmx19000M
```

Po týchto krokoch je možné zo zdrojového kódu v Jave zavolať nasledovné statické metódy, realizujúce proces trénovania resp. predikcie (estimácie):

Statická metóda pre realizáciu trénovania:

```
int fve.FVE_train.exec(String s1, String s2, String s3);
```

Statická metóda pre realizáciu predikcie (estimácie):

```
int fve.FVE_predict.exec(String r1, String r2, String r3, String r4);
```

Návratovými hodnotami oboch metód sú celé čísla typu int, ktoré indikujú návratový kód prípadnej chyby. Návratový kód 0 indikuje korektný stav bez chyby. Ostatné kódy chybových stavov, ako aj ich popis, resp. spôsob ich odstránenia sú uvedené používateľskej príručke v časti „Prehľad chybových kódov a spôsoby odstránenia chýb“.

Uvedené funkcie majú nasledovné vstupné parametre v tvare reťazcov:

s1 - plný názov súboru s príponou csv, s úplnou cestou, obsahujúci FVE dáta na trénovanie

s2 - plný názov súboru s príponou csv, s úplnou cestou, obsahujúci meteo-dáta na trénovanie

s3 – názov súboru - modelu s príponou .model, do ktorého bude natrénovaný model uložený.

r1 – plný názov súboru s príponou csv, s úplnou cestou, obsahujúci FVE dáta k estimácii

r2 – plný názov súboru s príponou csv, s úplnou cestou, obsahujúci meteo-dáta k estimácii

r3 – meno súboru modelu, s príponou model, ktorý bude použitý pre predikciu (estimáciu).

r4 – meno výstupného súboru s príponou csv, do ktorého budú zapísané výsledné odhadnuté hodnoty FVE.

poznámka: V prípade ak cesta k príslušnému súboru obsahuje medzeru, alebo iný špeciálny znak, je vhodné celú túto cestu uviesť do úvodzoviek "...". Cestu tiež doporučujeme oddeľovať tradičnými znakmi / , a teda vyhnúť sa znaku \ (backslash).

Volanie týchto metód môže byť napríklad nasledovné:

```
int err1 = fve.FVE_train.exec("D:/data/FVE_Data2021.csv",  
"D:/data/meteo_Data2021+2022.csv", "model1.model");  
  
int err2 = fve.FVE_predict.exec("D:/data/FVE_Data2022.csv",  
"D:/data/meteo_Data2021+2022.csv", "model1.model", "vystupy.csv");
```

Uvedený kód realizuje natrénovanie modelu do súboru model1.model, na základe FVE dát za rok 2021 a spoločných meteo-dát; Druhý riadok realizuje následnú predikciu na FVE údajoch z roku 2022.

Možné je tiež volanie statickej metódy main(), a to ako pre proces tréningovania,

```
public static void fve.FVE_train.main(String args []);
```

tak aj pre proces predikcie (estimácie):

```
public static void fve.FVE_predict.main(String args []);
```

V týchto metódach sú vstupné parametre umiestnené v poli reťazcov, pričom poradie vstupov je identické ako v prípade funkcií exec().

## 1.4 Štruktúra súborov

Program pre svoju činnosť používa súbory typu csv (comma separated values), ktoré sú plne kompatibilné, ľahko prenositeľné medzi systémami, tak aj ľahko čitateľné a editovateľné v ľubovoľnom textovom editore. Formát csv ako oddeľovač hodnôt používa znak čiarky. Pri reálnych číslach - ako desatinný oddeľovač (na oddelenie celej časti od desatinnej časti reálneho čísla) sa používa znak bodka. Konkrétne štruktúry jednotlivých súborov, ako aj významy jednotlivých hodnôt sú popísané opísané v jednotlivých podkapitolách.

### 1.4.1 Konfiguračný súbor settings.ini

Konfiguračný súbor settings.ini slúži na nastavenie ciest k jednotlivým dôležitým adresárom, ku ktorým program potrebuje pristupovať. Samotný súbor settings.ini musí byť umiestnený v rovnakom adresári ako hlavná knižnica FVE.jar, odkiaľ bude program spúšťaný.

Tento konfiguračný súbor obsahuje 3 riadky s úplnými cestami k špecifickým adresárom:

- Prvý riadok určuje plnú cestu k adresáru, kde môže estimačný systém vytvárať svoje dočasné pracovné súbory (čiže daný používateľ systému musí mať práva na zápis do tohto adresára; zároveň daný adresár musí byť umiestnený, na zariadení, kde je možné realizovať priamo zápis (teda nie DVD a pod.) a taktiež kde je dostatočný voľný diskový priestor).
- Druhý riadok určuje plnú cestu k adresáru obsahujúcemu potrebné súbory číselníkov. Patria sem súbory číselníkov: ciselnik\_instal\_vykonu.csv a lokality\_FVE\_ciselnik.csv Štruktúra jednotlivých číselníkov je popísaná v časti 1.4, a musí byť korektná.
- Tretí riadok určuje plnú cestu k adresáru, určenému pre výstupy. Do tohto adresára doporučujeme umiestňovať súbory natrénovaných modelov ako aj výstupy estimácie.

Na oddelovanie adresárov doporučujeme použiť bežné lomítka ( znak '/' ), ktorý funguje ako v operačnom systéme Linux, tak aj Windows. Taktiež je potrebné rozlišovať veľké a malé písmená v názvoch adresárov.

Príklad vyplnenia súboru settings.ini:

```
c:/Users/Peter/Documents/NetBeansProjects/FVE/tmpDir/  
c:/Users/Peter/Documents/NetBeansProjects/FVE/ciselniky/  
c:/Users/Peter/Documents/NetBeansProjects/FVE/vystupy/
```

## 1.4.2 Vstupné dátové súbory

### Súbor s FVE dátami

Tento vstupný dátový súbor má formát CSV, pričom jednotlivé hodnoty v riadku sú oddelené čiarkami. Prvý riadok súboru predstavuje hlavičku súboru s menami jednotlivých stĺpcov; Nasledujúce riadky súboru zodpovedajú odlišným časovým okamžikom, resp. dátumom s príslušnými hodnotami FVE uvedených v MW.

Prvý stĺpec súboru má názov DateTime, a reprezentuje dátum a čas (vo formáte dd.MM.yyyy hh:mm), ktorý sa viaže pre daný riadok. Ostatné stĺpce reprezentujú jednotlivé FVE lokality (vo formáte Loc01, Loc02, ...);

Jednotlivé číselné hodnoty v tabuľke predstavujú meraný výkon FVE (v MW) pre daný dátum (určený riadkom) a pre danú lokalitu (určenú príslušným stĺpcom). V prípade neznámej hodnoty meraného výkonu, sa neuvádza žiadna hodnota, a teda v súbore sú uvedené 2 čiarky bezprostredne za sebou.

### Príklad štruktúry súboru s FVE dátami pre 3 FVE Lokality (Loc01 – Loc03)

```
DateTime,Loc01,Loc02,Loc03
17.01.2019 05:00,0.0,0.0001,0.0
17.01.2019 13:00,11.48,8.59,23.17
17.01.2019 13:30,11.97,9.41,31.87
```

### Súbor s meteo-dátami

Súbor s meteo-dátami má výrazne podobnú štruktúru ako súbor s FVE dátami. Prvý riadok súboru opäť tvorí hlavička súboru; ostatné riadky predstavujú odlišné časové momenty (vo formáte dd.MM.yyyy hh:mm).

Jednotlivé stĺpce reprezentujú 3 odlišné meteo-veličiny (žiarenie, vlhkosť a teplotu), pričom každý z nich je párovaný so všetkými dostupnými meteo-lokalitami (ML01, ML02, ...).

Úvodný stĺpec s názvom DateTime popisuje informáciu o dátume a čase (vo formáte dd.MM.yyyy hh:mm). Ostatné stĺpce sú kombináciou mena veličiny a identifikátora príslušnej meteo-lokality. Napr. stĺpec s názvom „žiarenie\_ML04“ reprezentuje veličinu žiarenia pre meteo-lokalitu ML04. Hodnoty v tomto stĺpci teda sú reálne čísla, predstavujúce intenzitu žiarenia v meracej lokalite ML04, pričom príslušný dátum a čas nameranej intenzite žiarenia určuje riadok v tabuľke.

### Príklad štruktúry súboru s meteo-dátami pre 2 meteo-lokality

```
DateTime,ziarenie_ML01,ziarenie_ML02,teplota_ML01,teplota_ML02,vlhkost_01,vlhkost_02
17.01.2019 05:00,0.0,0.0,-6.5,-3.2,48.1,65.2
17.01.2019 13:00,631.1,842.4,-1.5,1.1,49.5,57.7
17.01.2019 13:30,658.4,847.3,-1.4,1.7,49.4,56.9
```

### 1.4.3 Súbory číselníkov

Všetky súbory číselníkov sú vo formáte CSV súborov (comma separated values), s oddeľovačom čiarkou. Ako desatinný oddeľovač pri reálnych číslach (na oddelenie celej od desatinnej časti) je použitý znak bodky. Adresár číselníkov (určený v súbore settings.ini) musí obsahovať 2 číselníky a to konkrétne súbory: `ciselnik_instal_vykonu.csv` a `lokality_FVE_ciselnik.csv`

**Súbor `ciselnik_instal_vykonu.csv`** obsahuje informácie o inštalovanom výkone (meranom aj nemeranom) pre všetky používané lokality pre stanovené dátumy merania. Musí taktiež obsahovať odhad inštalovaného výkonu do budúcnosti (ten môže byť odhadnutý napr. na základe predchádzajúceho ročného percentuálneho nárastu).

Tento súbor obsahuje vždy 4 stĺpce určené názvami:  
`Date,Location,InstVyk_merany,InstVyk_nemerany`

Tie sú zároveň uvedené v prvom riadku súboru, oddelené čiarkami. Ostatné riadky obsahujú dáta, pričom každý riadok reprezentuje jeden záznam o určení inštalovaného meraného a nemeraného výkonu pre danú lokalitu a dátum.

Významy jednotlivých stĺpcov v poradí v akom sú v súbore sú nasledovné:

`Date` – dátum (v tvare deň.mesiac.rok), pre ktorý sú hodnoty výkonu aktuálne (v prípade ak nie je striktné dané, k akému dátumu sú hodnoty aktuálne, je možné použiť napr. deň v strede mesiaca a pod.).

`Location` – lokalita (tvare `Loc01` až `Loc99`) pre ktorú je stanovený inštalovaný výkon

`InstVyk_merany` – hodnota určeného/odhadnutého inštalovaného výkonu (meraného) v MW

`InstVyk_nemerany` - hodnota určeného/odhadnutého inštalovaného výkonu (nemeraného) v MW

Príklad súboru `ciselnik_instal_vykonu.csv` pre 3 lokality (`Loc01` až `Loc03`) je nasledovný (uvedené hodnoty nezodpovedajú skutočným hodnotám inštalovaného výkonu):

```
Date,Location,InstVyk_merany,InstVyk_nemerany
1.1.2019,Loc01,10.5,1.5
1.1.2019,Loc02,40,4.5
1.1.2019,Loc03,21.3,2.2
15.6.2020,Loc01,12.1,1.8
12.6.2020,Loc02,45,5.5
12.6.2020,Loc03,23.5,2.4
21.2.2021,Loc01,13.4,2.1
18.2.2021,Loc02,50.3,5.9
16.6.2021,Loc03,25.1,2.6
1.1.2026,Loc01,17.8,3.0
1.1.2026,Loc02,61.1,7.2
17.1.2026,Loc03,30.7,3.1
```



### **Súbor lokality\_FVE\_ciselnik.csv**

Tento súbor obsahuje 6 stĺpcov; Počet riadkov s údajmi zodpovedá počtu použitých FVE lokalít (určených ako Loc01 až Loc99, najviac 99 lokalít). Prvý riadok reprezentuje hlavičku súboru. Súbor obsahuje nasledovné stĺpce oddelené čiarkami (zodpovedá hlavičke súboru): FVELocationID,FVELocationName,NearMeteo1,NearMeteo2,NearMeteo3,NearMeteo4

#### Významy jednotlivých stĺpcov v súbore sú nasledovné:

FVELocationID – Identifikátor danej lokality v tvare Loc01 až Loc99

FVELocationName – Textové meno lokality

NearMeteo1 – Príslušný identifikátor najbližšej/najrelevantnejšej meteo-Lokality (v tvare ML01 – ML99), ktorej meteo-dáta budú použité pri modelovaní.

NearMeteo2 – Príslušný identifikátor 2. najbližšej/najrelevantnejšej meteo-Lokality (v tvare ML01 – ML99), ktorej meteo-dáta budú použité pri modelovaní.

NearMeteo3 – Príslušný identifikátor 3. najbližšej/najrelevantnejšej meteo-Lokality (v tvare ML01 – ML99), ktorej meteo-dáta budú použité pri modelovaní.

NearMeteo4 - Príslušný identifikátor 4. najbližšej/najrelevantnejšej meteo-Lokality (v tvare ML01 – ML99), ktorej meteo-dáta budú použité pri modelovaní.

Priradenie jednotlivých Meteo-lokalít ku Lokalitám FVE môže byť realizované na základe geografickej vzdialenosti, či maximálnej korelácie produkcie FVE voči miere žiarenia v danej lokalite.

#### **Príklad súboru lokality\_FVE\_ciselnik.csv (demonštratívny príklad) :**

```
FVELocationID,FVELocationName,NearMeteo1,NearMeteo2,NearMeteo3,NearMeteo4  
Loc01,SSE – Zapadna Lokalita,ML01,ML02,ML03,ML05  
Loc02,SSE – Stredna Lokalita,ML02,ML03,ML05,ML06  
Loc03,SSE – Vychodna Lokalita,ML05,ML06,ML03,ML02
```

#### 1.4.4 Výstupné súbory

Medzi výstupné súbory zaraďujeme súbory modelov (ako výstup tréningovej časti), ako aj výstup estimačnej časti – výstupný csv súbor obsahujúci estimované hodnoty. Všetky tieto výstupné súbory odporúčame ukladať do výstupného adresára (špecifikovaného v konfiguračnom súbore settings.ini ).

##### Súbory modelov

Štruktúra súboru modelu je netriviálna, nakoľko sa jedná o binárny súbor serializujúci Javovské objekty rozličných inštancií. Model taktiež obsahuje číselníky, ktoré boli použité pri procese tréningovania a to z dôvodu zabezpečenia korektnej dátovej integrity. Preto nedoporučujeme zasahovať do obsahu týchto súborov a žiadnym spôsobom v nich vykonávať úpravy. Jediným korektným spôsobom úprav je vytvorenie (natréningovanie) nového modelu použitím korektného procesu tréningovania daného programu, nad korektnými dátami.

Modely, ktoré dosahujú uspokojivé hodnoty estimácie je vhodné zarchivovať, pre možné neskoršie použitie (napr. napáliť na DVD, či skopírovať na USB stick). Súbory modelu sú relatívne rozsiahle (model obsahujúci 16 lokalít zaberá cca 600 MB). Preto pri archivácii môže byť použitá aj bezstratová kompresia (napr. zip), ktorá do určitej miery rozsah zredukuje.

##### Výstupný súbor estimácií typu CSV

Výstupom estimačnej časti je súbor formátu csv. Jednotlivé hodnoty sú teda oddelené čiarkou, a desatinným oddeľovačom je znak bodka. Jednotlivé riadky v súbore zodpovedajú jednotlivým dátovým záznamom, pričom prvý riadok definuje hlavičku. Jednotlivé stĺpce predstavujú atribúty rovnakého typu.

Tento výstupný súbor má 5 stĺpcov, pomenovaných nasledovne (jednotlivé názvy sú oddelené opäť čiarkami): Location,DateTime,realFVE,predictedFVE,predictedFVENonMeas

Významy jednotlivých stĺpcov - v danom poradí, sú nasledovné:

**Location** – lokalita, pre ktorú sa modeluje výkon FVE (Loc01 až Loc99)

**DateTime** – dátum a čas, pre ktorý sa modeluje výkon FVE (vo formáte dd.MM.yyyy hh:mm)

**realFVE** – skutočná hodnota výkonu FVE, ktorá bola dodaná vo vstupom súbore FVE dát k procesu estimácie (desatinné číslo). V prípade ak táto hodnota nebola zadaná vo vstupnom súbore FVEdata.csv (súbor obsahoval chýbajúce hodnoty), tak aj tento stĺpec obsahuje chýbajúce hodnoty. Hodnota je v MW.

**predictedFVE** – estimovaná hodnota meraného výkonu FVE, odhadnutá modelom pre príslušný dátum, čas a lokalitu (desatinné číslo). Hodnota je v MW.

**predictedFVENonMeas** - estimovaná hodnota nameraného výkonu FVE (desatinné číslo v MW), určená výpočtom:

$\text{predictedFVENonMeas} = \text{predictedFVE} * \text{NonMeasRatio};$

kde hodnota NonMeasRatio určuje podiel hodnôt nameraného k meranému výkonu FVE na základe číselníka inštalovaných výkonov pre daný deň a danú lokalitu.

### Príklad štruktúry výstupného súboru

(nezodpovedá reálnym lokalitám ani údajom, slúži len na demonštráciu):

```
Location,DateTime,realFVE,predictedFVE,predictedFVENonMeas
Loc04,13.03.2017 07:15,0.209707,0.35982989,0.03586997
Loc05,13.03.2017 07:15,0.052466,0.06780809,0.00678081
Loc06,13.03.2017 07:15,0.0,0.00702588,7.0259E-4
Loc07,13.03.2017 07:15,9.738296,8.89296053,0.89085896
Loc08,13.03.2017 07:15,0.619961,0.54976077,0.05351653
Loc09,13.03.2017 07:15,0.790887,0.6906374,0.06919187
Loc10,13.03.2017 07:15,3.777837,2.39201771,0.24092265
Loc11,13.03.2017 07:15,0.0,5.6899E-4,5.69E-5
Loc12,13.03.2017 07:15,0.715117,0.49831795,0.04941996
Loc13,13.03.2017 07:15,1.888945,1.7529461,0.17655572
Loc14,13.03.2017 07:15,2.912704,2.62289162,0.26060782
Loc15,13.03.2017 07:15,0.258166,0.2236637,0.02324191
Loc16,13.03.2017 07:15,2.913378,2.39541612,0.24167403
Loc17,13.03.2017 07:15,0.40193,0.45273451,0.04527345
```

#### 1.4.5 Previazanie dát

Jednotlivé dátové súbory musia vzájomne korešpondovať; nemôžu teda obsahovať vzájomne si odporujúce údaje. Musia spĺňať určité pravidlá pre zachovanie korektnej dátovej integrity:

- Všetky lokality uvedené v súbore FVE dát, musia byť obsiahnuté v číselníku lokality\_FVE\_ciselnik.csv ako samostatný riadok
- Všetky meteo-lokalitty (napr. ML01) vyskytujúce sa v číselníku lokality\_FVE\_ciselnik.csv musia byť zastúpené aj v meteo-údajoch ako samostatné stĺpce, a to v kombinácií pre všetky 3 veličiny (ziarenie, vlhkosť aj teplota).
- Všetky dátumy zaradené v dátovom súbore FVE musia byť v súbore **ciselnik\_instal\_vykonu.csv** dátumovo ohraničené zdola aj zhora a to pre všetky FVE lokality zastúpené v súbore FVE údajov. Inak povedané, ak sa napr. v dátovom FVE súbore nachádza záznam z dátumu 1.6.2018, pre lokalitu Loc03, tak potom nutne v číselníku inštalovaných výkonov **ciselnik\_instal\_vykonu.csv**, musí byť aspoň 1 riadok pre lokalitu Loc03 s dátumom pred týmto termínom (1.6.2018), a aspoň 1 s dátumom po tomto termíne.

## 1.5 Prehľad chybových kódov a spôsoby odstránenia chýb

Tabuľka chybových kódov s popisom chyby

ErrorCode	Popis chyby
0	Stav bez chyby
-4	Chyba pri manipulácii so súborom settings.ini (súbor nebol nájdený, alebo má nesprávnu štruktúru / je poškodený).
-5	Chyba pri uzatváraní súboru settings.ini
-20	Neúspešný zápis do súboru
-31	V číselníku inštalovaných výkonov je nekorektný dátum
-32	V číselníku inštalovaných výkonov je nekorektná hodnota (nečíselná hodnota / Inf / NaN)
-33	V číselníku inštalovaných výkonov je nekorektná lokalita (musí mať 5 znakov)
-34	V číselníku inštalovaných výkonov je nekorektná lokalita (musí začínať na "Loc")
-35	Lokalita v číselníku inštalovaných výkonov musí končiť na 2 cifry
-40	Chyba pri zápise dočasného súboru číselník_instal_výkonu2.csv
-41	Chyba pri zápise dočasného súboru lokality_FVE_číselník.csv
-42	Chyba pri načítaní modelu (súbor modelu je nenájdený, alebo poškodený / nekorektný)
-44	Chyba pri čítaní dočasného súboru tmp3.csv
-45	Chyba pri čítaní dočasného súboru tmp3p.csv
-48	Chyba pri čítaní csv súboru
-50	Chyba pri trénovaní modelu
-74	Chyba, lokalita nenájdená v číselníku
-75	Chyba, lokalita nenájdená v údajoch FVE
-76	FVE DateTime atribút nenájdený
-77	Atribút DateTime nebol nájdený v meteo údajoch.
-78	Chyba pri dátovej integrácii vo fáze 1.
-104	Nekorektná štruktúra csv dát na estimáciu
-108	Chyba pri zápise csv súboru na disk
-111	Chyba pri dátovej integrácii vo fáze 2.
-120	Chyba pri výpočte presnosti modelu: Pole nebolo alokované;
-999	Chyba pri serializácii modelu do súboru

Tabuľka chybových kódov s uvedením možností nápravy chyby:

ErrorCode	Možné príčiny / spôsob odstránenia chyby
-4	Overte korektnosť názvu a umiestnenia súboru settings.ini Overte korektnosť nastavení v súbore settings.ini, a práva na čítanie tohto súboru
-5	Overte, či súbor settings.ini nemáte otvorený na zápis
-20	Overte, či je na disku dostatočný voľný priestor (aspon 4GB) Overte, či prihlásený používateľ má právo zapisovať do adresára dočasných súborov ako aj adresára výstupov. Overte, či nejaký zo súborov (dátových, číselníkov, alebo výstupov) nemáte už otvorený iným programom na zápis.
-31	Opravte údaje o dátume v číselníku inštalovaných výkonov.
-32	Opravte údaje v číselníku inštalovaných výkonov.
-33	Opravte údaje o lokalite v číselníku inštalovaných výkonov.

-34	Opravte údaje o lokalite v číselníku inštalovaných výkonov.
-35	Opravte údaje o lokalite v číselníku inštalovaných výkonov.
-40	Overte, či prihlásený používateľ má právo zapisovať do adresára dočasných súborov. Overte, či je na disku dostatočný voľný priestor (aspon 4GB)
-41	Overte, či prihlásený používateľ má právo zapisovať do adresára dočasných súborov. Overte, či je na disku dostatočný voľný priestor (aspon 4GB)
-42	Overte, či je parameter - meno modelu uvedený správne a či sa nachádza v správnom adresáry Overte, či tento súbor nie je otvorený inou aplikáciou. V prípade poškodenia súboru s modelom, je vhodné jeho obnovenie zo zálohy
-44	Overte, či žiadny program nepracuje s týmto súborom Overte správnosť nastavenia adresára pre dočasné súbory v súbore settings.ini
-45	Overte, či žiadny program nepracuje s týmto súborom Overte správnosť nastavenia adresára pre dočasné súbory v súbore settings.ini
-48	Overte, či vstupné csv súbory majú špecifikovanú štruktúru Overte, či používateľ má práva na čítanie vstupných csv súborov
-50	Overte, či vstupné údaje zodpovedajú predpísanej štruktúre
-74	Doplňte lokality použité vo FVE-dátach, do číselníkov
-75	Doplňte lokality použité vo FVE-dátach, do Meteouúdajov
-76	Doplňte do dátového súboru stĺpec DateTime
-77	Opravte údaje v súbore meteo údajov
-78	Overte, či vstupné dátové súbory majú špecifikovanú štruktúru Overte, či súbory číselníkov majú špecifikovanú štruktúru Overte, či je na disku dostatočný voľný priestor (aspon 4GB) V prípade ak počas behu programu sa množstvo voľnej pamäte ram extrémne znížilo, tak navýšte hodnotu pri spúšťaní programu (parameter -Xmx)
-104	Opravte štruktúru dát na estimáciu podľa špecifikácie
-108	Overte, či je na disku dostatočný voľný priestor (aspon 4GB) Overte, či prihlásený používateľ má právo zapisovať do adresára dočasných súborov ako aj adresára výstupov. Overte, či nejaký zo súborov (dátových, číselníkov, alebo výstupov) nemáte už otvorený iným programom na zápis.
-111	Overte, či je na disku dostatočný voľný priestor (aspon 4GB) Overte, či majú vstupné dáta a číselníky správnu štruktúru V prípade ak počas behu programu sa množstvo voľnej pamäte ram extrémne znížilo, tak navýšte hodnotu pri spúšťaní programu (parameter -Xmx)
-120	Overte, či dáta na predikciu obsahujú záznamy (riadky) s platnými hodnotami Overte si, či máte dostatok voľnej pamäte (windows: taskmanager, linux: free -m) Zvýšte objem pridelennej pamäte pri spúšťaní programu (príkazom pri spúšťaní -Xmx)
-999	Overte, či je na disku dostatočný voľný priestor (aspon 4GB) Overte, či má používateľ právo zapisovať do adresára s výstupmi Overte, či je cesta pre výstupný adresár v súbore settings.ini nastavená správne Overte, či nejaký zo súborov (dátových, číselníkov, alebo výstupov) nemáte už otvorený iným programom na zápis.