Clustering of pharmacokinetic responses using model selection based on MDL or NML

Rui Guerra, Paulo Mateus and Alexandra M. Carvalho

App user guide

App setting up and run

The pkclusteringmdl.jar app is written in java so you need to have java runtime environment (JRE) installed in your machine (version 8 at least). If you are not sure about the version you have installed, please type on the command line:

```
java –version
```

If you need to install another version, please type on your search engine “download java” and follow the instructions of the site.

You then need to download the app pkclusteringmdl.jar from:

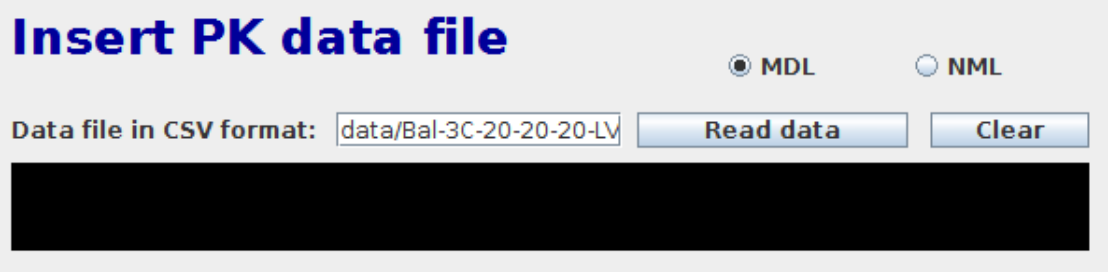
```
https://rjri.github.io/pkclusteringmdl/
```

You can place it anywhere you like. To run the app open a command line and type:

```
java –jar pkclusteringmdl.jar
```

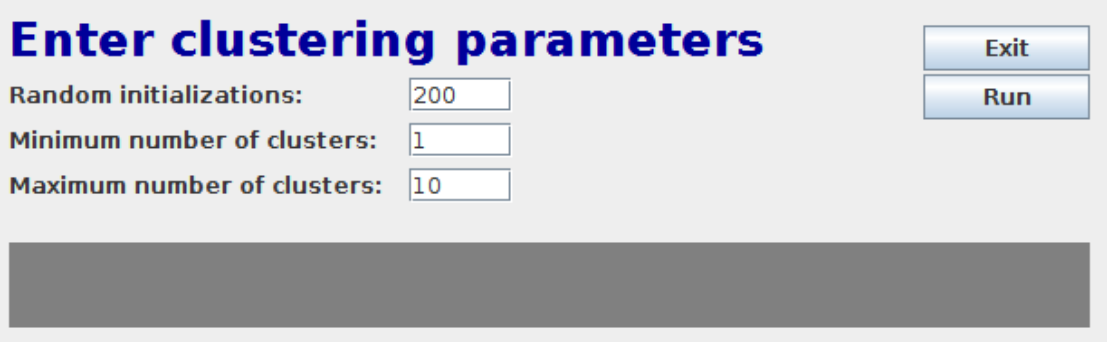
Running the app

When you run the application the follow window appears:

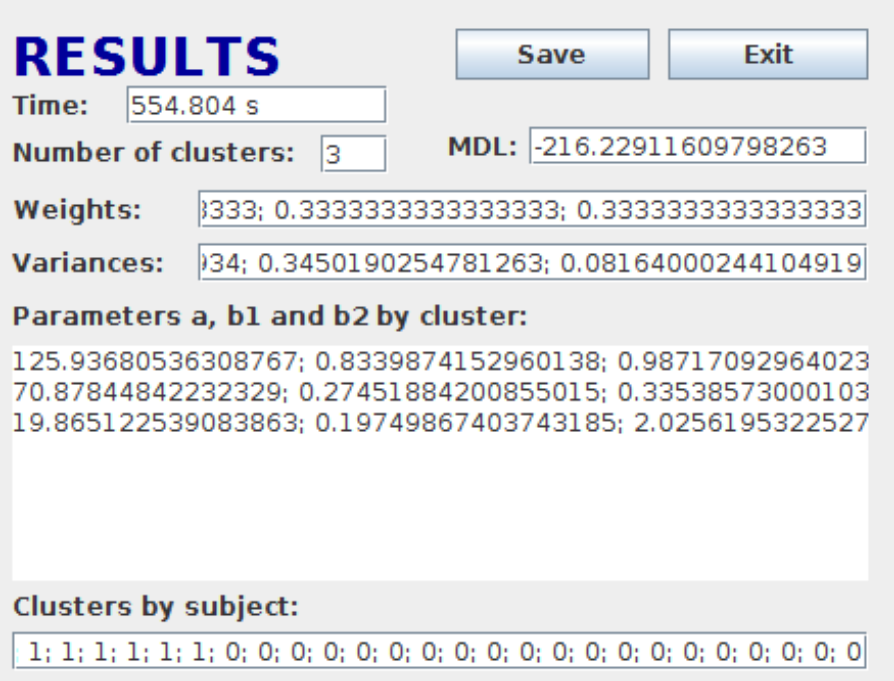


Please insert the name of the data you want to process (full name with extension); you can use, for instance, data/Bal-3C-20-20-20-LV.csv from the data folder offered in the web page along with the application. Note that this file must be in the same directory as the executable jar file; alternatively, you may pass a complete path to the file. In addition, you may choose the model selection criterion to be used in the execution of the program, MDL or NML, by selecting the corresponding button. After that press the “Read” button.

A new window appears where you have to indicate the number of random initializations and the minimum and maximum number of clusters the program may output. For instance, you may choose 200 random initializations and choose a minimum of 1 and a maximum of 10 clusters. Then press “Run”.



After pressing the “Run” button you need to wait for the results, which will appear in another window:



This is a window with the results. In:

- “Time” – Total execution time.
- “Number of clusters” – optimal number of clusters.
- “MDL” or "NML" – the function used to evaluate the goodness of the solution, known as stochastic complexity.
- “Weights” – weight of each cluster separated by a semicolon.
- “Variances” – variance of each cluster separated by a semicolon.
- “Parameters a, b1 and b2 by cluster” – the parameters (α, β_1 and β_2) of each of the retrieved clusters; each cluster in a line, with parameters separated by a semicolon.

- “Clusters by subject” – indicates the cluster of each subject.

You can also save the results by pressing in the “Save” button to a file (automatically) named Results.txt (that will appear in the same folder as the app).

In the end, please, close all windows.