

2021

# SelectWave User Manual

DEVELOPERS

DR. FATİH KAHRIMAN, DR. KRISTIAN HOVDE LILAND

BAF ELECTRONIC SOFTWARE AGRICULTURE INC. | Çanakkale Tecnoogy Development Zone, Turkey

## Preface

Multivariate modeling analysis is nowadays used in agriculture, health, chemistry and many other areas. Spectral data analysis is also one of the analysis methods that has become widespread in scientific studies. In this analysis method, which is based on multivariate modeling, creating calibration models by associating the data obtained from different devices with the target variable is one of the main objectives.

There are many software for multivariate data analysis that require installation and are expensive. However, it is not possible to access these programs in most scientific research and student studies. Although these analyzes can be performed on some free platforms (such as R), there are various obstacles to the user such as installation and code writing on these platforms. Therefore, there is a need for tools that can perform data analysis practically. In this context, it is thought that web applications, which have become quite popular in recent years and can be accessed free of charge on the internet, can meet this need.

In this context, the SelectWave application, which can be accessed on the web, is free, user-friendly and offers comprehensive data analysis solutions. This manual has been prepared in order to share the practical use of the SelectWave application and the points that need to be paid attention to by users. We hope that the SelectWave application will be useful to academics, students and other researchers conducting multivariate modeling studies.

Dr. Fatih Kahrıman

Dr. Kristian Hovde Liland

## Content

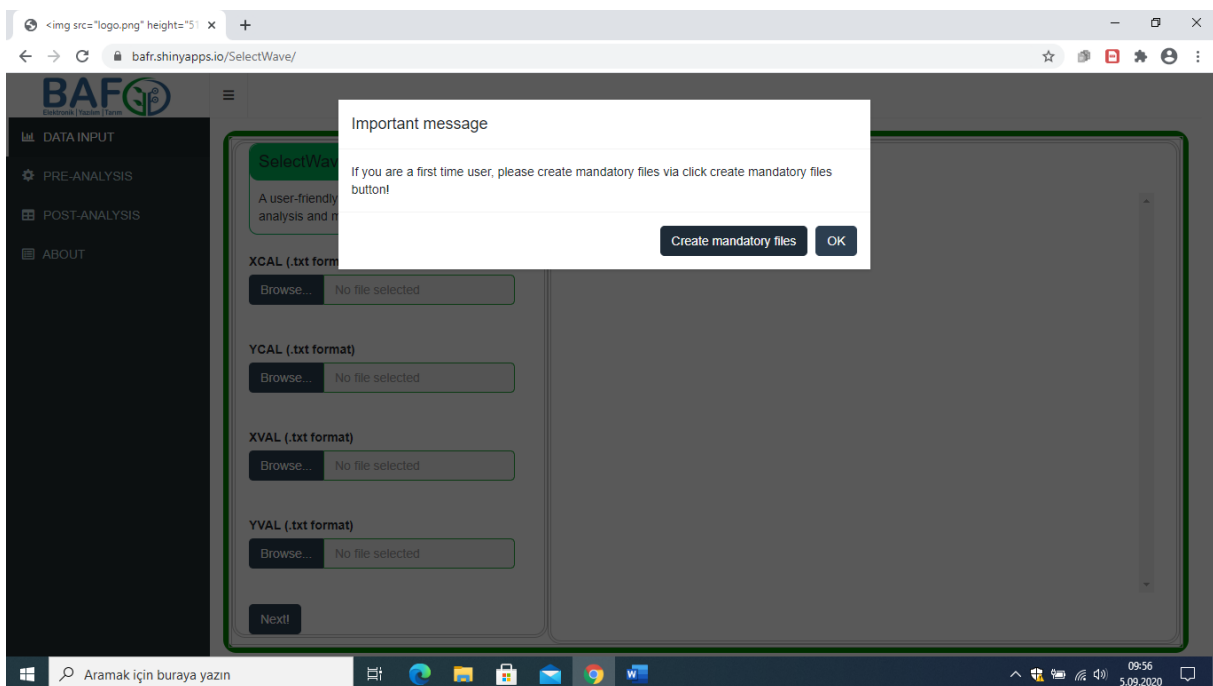
|                          |    |
|--------------------------|----|
| Preface .....            | 1  |
| Opening SelectWave ..... | 3  |
| Data Loading.....        | 4  |
| Pre-Analysis.....        | 6  |
| Post-Analysis .....      | 10 |
| About .....              | 16 |

## Opening SelectWave

The SelectWave application can be accessed from the link below on the internet.

<https://bafr.shinyapps.io/SelectWave/>

The user encounters a warning message when the application is opened for the first time. As stated in this message, if the SelectWave application is to be used for the first time, the button "Create mandatory files" is clicked. Then, the analysis continues with the "OK" button. If you have used this application before, you can directly click the "OK" button. The mandatory files button is required for saving the analysis results as csv to the user's workspace and for the application to run smoothly.



After this approval process is completed, the first step for analysis; "DATA INPUT" menu will become available.

**Important: First time users of the SelectWave application must create the necessary files with the "Create mandatory files" button.**

## Data Loading

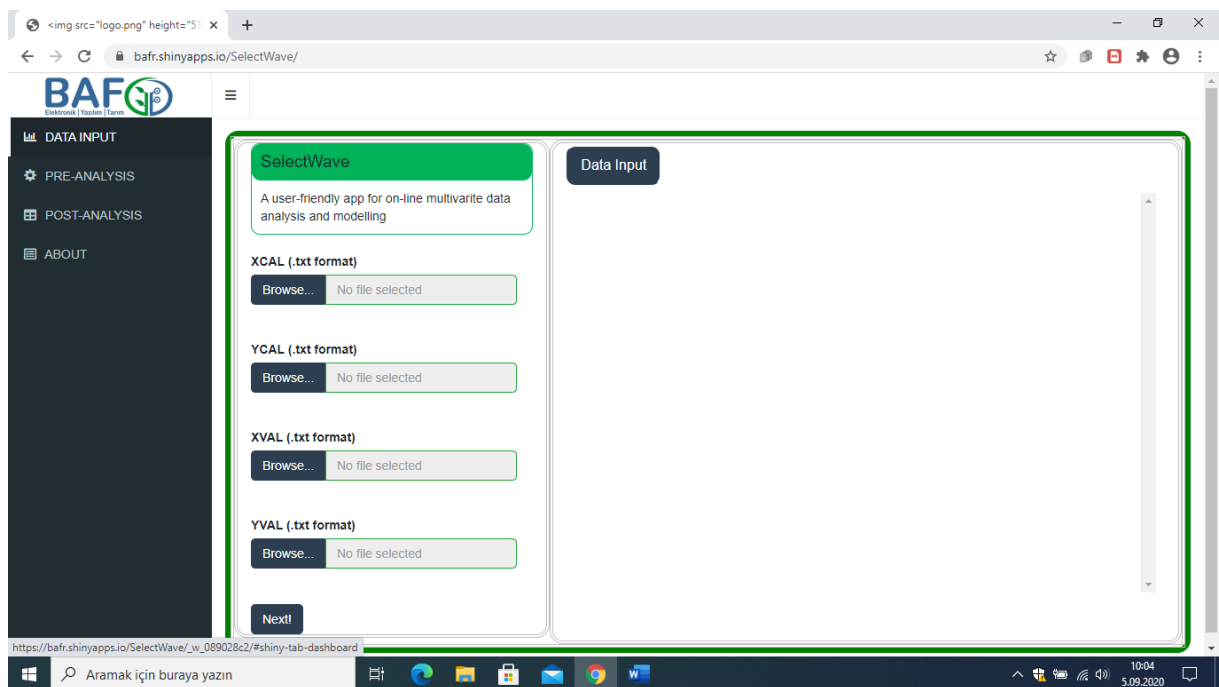
To load data into the SelectWave application, the application's DATA INPUT menu opens. Four separate data entries must be made for model development and validation. From the four file upload menus in the interface;

**XCAL:** Data to be used as predictive variables in the calibration model to be created (must be loaded).

**YCAL:** In the prediction model to be created, data regarding the quantitative dependent variable (must be loaded).

**XVAL:** Data regarding the quantitative dependent variable in the validation model (must be loaded).

**YVAL:** In order to verify the created estimation model, data related to the dependent variable to be used must be loaded, otherwise cross-validation is used instead.



Important: All files to be uploaded to the SelectWave application must have a .txt extension. In order to prepare these files, first the data to be used in the analysis should be imported into excel file. Then this data should be transferred to a separate notepad file using the copy-paste method. File naming has no effect on the analysis result.

Spectral data in the sample data set of the application are recorded at 1 nm intervals between 1200-2400 nm. The XCAL file contains spectral data of 150 different samples, and the XVAL file contains spectral data of 50 samples. The YCAL file contains amylopectin content of 150 samples in amylopectin spectral data, and the YVAL file contains data on amylopectin content of 50 samples for external validation. When the entry of these data is completed, the first five lines of the spectral data in the calibration set are displayed on the user screen.

The screenshot shows the SelectWave application interface. On the left is a dark sidebar with a menu: DATA INPUT (selected), PRE-ANALYSIS, POST-ANALYSIS, and ABOUT. The main area is titled 'SelectWave' and contains a description: 'A user-friendly app for on-line multivariate data analysis and modelling'. Below this are four sections for file uploads: XCAL (.txt format), YCAL (.txt format), XVAL (.txt format), and YVAL (.txt format). Each section has a 'Browse...' button and an 'Upload complete' button. At the bottom of these sections is a 'Next!' button. To the right of the upload sections is a 'Data Input' table with 10 columns (X1200 to X1209) and 6 rows of data.

| X1200 | X1201 | X1202 | X1203 | X1204 | X1205 | X1206 | X1207 | X1208 | X1209 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 0.19  | 0.19  | 0.19  | 0.19  | 0.19  | 0.19  | 0.18  | 0.18  | 0.18  | 0.18  |
| 0.19  | 0.19  | 0.19  | 0.19  | 0.19  | 0.19  | 0.18  | 0.18  | 0.18  | 0.18  |
| 0.19  | 0.19  | 0.19  | 0.19  | 0.19  | 0.19  | 0.18  | 0.18  | 0.18  | 0.18  |
| 0.13  | 0.13  | 0.13  | 0.13  | 0.13  | 0.13  | 0.13  | 0.13  | 0.13  | 0.12  |
| 0.13  | 0.13  | 0.13  | 0.13  | 0.13  | 0.13  | 0.13  | 0.13  | 0.13  | 0.12  |
| 0.17  | 0.17  | 0.17  | 0.17  | 0.17  | 0.17  | 0.17  | 0.17  | 0.17  | 0.17  |

After completing the data entry, the "NEXT" button under the DATA INPUT menu is clicked to switch to the "PRE-ANALYSIS" menu.

## Pre-Analysis

In the "PRE-ANALYSIS" menu of the SelectWave application, a pre-model is created according to the PLS method as well as derivative and pre-treatment applications.

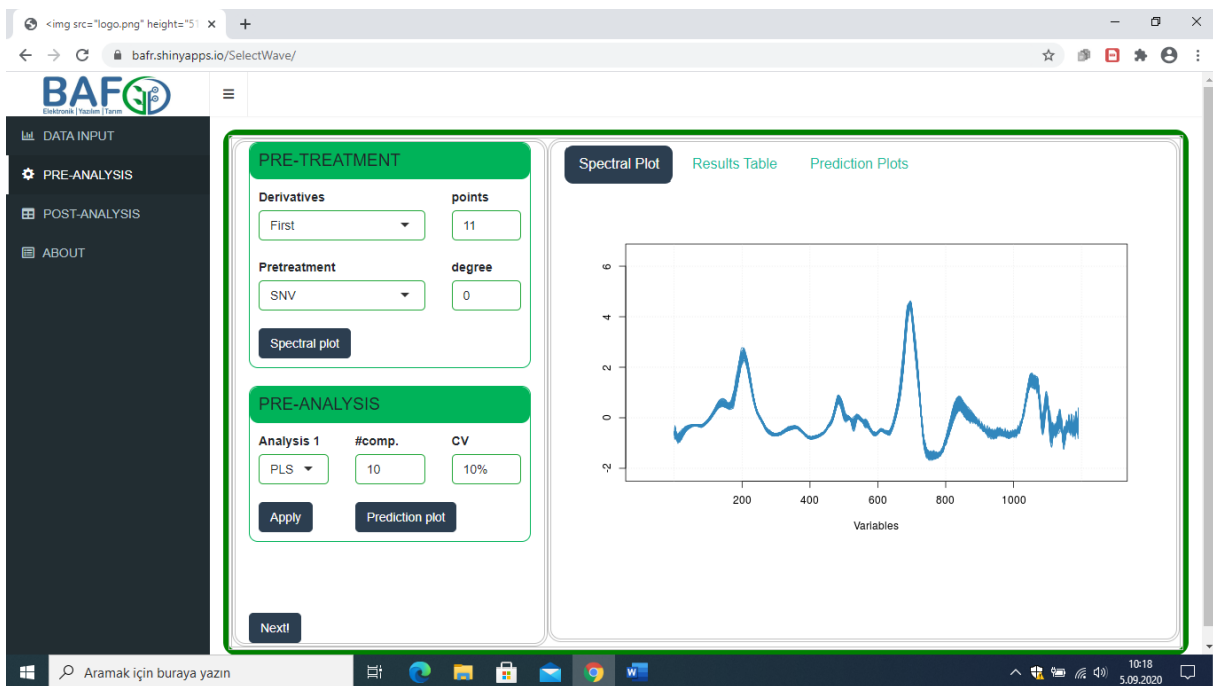
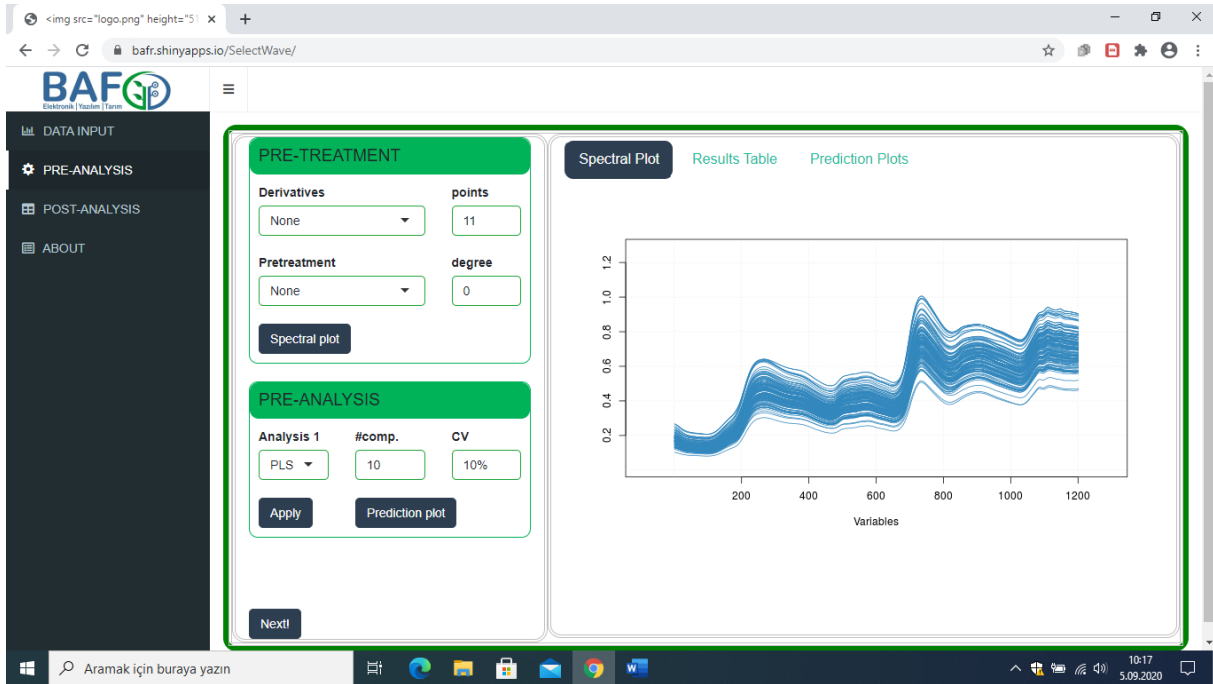
The screenshot displays the 'PRE-ANALYSIS' interface of the SelectWave application. On the left is a dark sidebar with menu items: 'DATA INPUT', 'PRE-ANALYSIS' (active), 'POST-ANALYSIS', and 'ABOUT'. The main workspace is divided into two primary sections: 'PRE-TREATMENT' and 'PRE-ANALYSIS'. The 'PRE-TREATMENT' section includes a 'Derivatives' dropdown menu set to 'None', a 'points' input field with the value 11, a 'Pretreatment' dropdown menu set to 'None', a 'degree' input field with the value 0, and a 'Spectral plot' button. The 'PRE-ANALYSIS' section features an 'Analysis 1' dropdown menu set to 'PLS', a '#comp.' input field with the value 10, a 'CV' input field with the value 10%, and buttons for 'Apply' and 'Prediction plot'. A 'Next!' button is located at the bottom left of the main workspace. On the right side of the workspace, there are three tabs: 'Spectral Plot' (active), 'Results Table', and 'Prediction Plots'.

The preprocesses that can be applied to the data uploaded to the application are listed below.

**Derivatives:** The options are None, First, Second, Smoothing. The window width point value defined for the derivatives is 11 and should be an odd number.

**Pretreatment:** None, SNV, MSC, EMSC, Baseline-polynomial options are available in the SelectWave application as data preprocessing. Data transformation is performed by the application according to the combination of pretreatment with the derivatives option selected by the user. The degree value is suggested by the software to match the selected pretreatment. The user can change these values with the desired value.

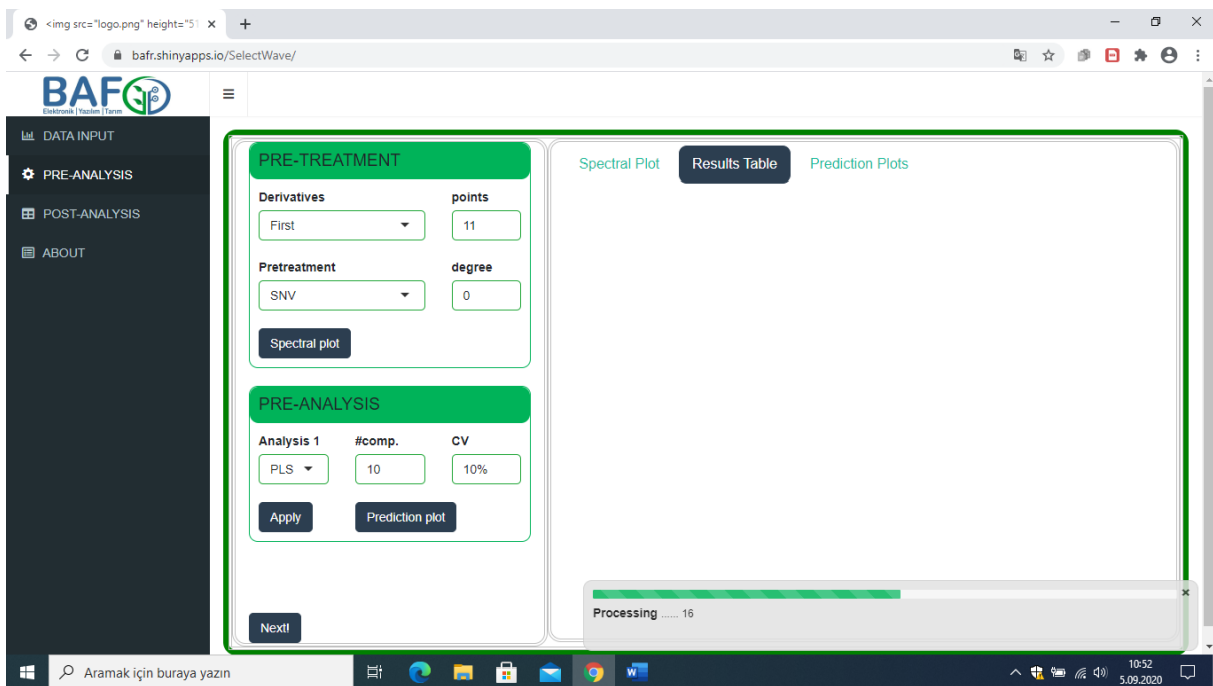
In order to see the effect of data transformation on spectral data, the "Spectral Plot" button is pressed. A spectral graph suitable for data transformation is displayed on the user screen. The screenshots below show the results without pretreatment (Derivatives: None, Pretreatment: None) and with pre-treatment options (Derivatives: FD, Pretreatment: SNV), respectively.





In the Pre-Analysis menu, click the "Apply" button to create a model with the PLS technique. In the example in this manual, the spectral pretreatment has been assigned as Derivatives: FD (points: 11) and Pretreatment: SNV (degree: 0). For the PLS analysis, the COMP value was defined as 10 and CV as 10%, and the screenshots of this analysis are presented below.

Important: The PLS technique is used in the preliminary analysis of the SelectWave application. In the PLS technique, the number of components can be assigned by the user and the CV value is assigned as "1" for LOO validation. If there is no separate data set for external validation, that is, if the calibration and external validation data consist of the same files, it is recommended that the CV value is not assigned 1.



A progress bar appears on the screen to complete the analysis. According to the analysis result, the calibration evaluation statistics can be seen in the "Results Table" sub-menu. These statistics can be downloaded as csv or excel files or copied to the clipboard and transferred into another file.

BAF

DATA INPUT

PRE-ANALYSIS

POST-ANALYSIS

ABOUT

**PRE-TREATMENT**

Derivatives: First (11 points)

Pretreatment: SNV (0 degree)

Spectral plot

**PRE-ANALYSIS**

Analysis 1: PLS (#comp.: 10, CV: 10%)

Apply Prediction plot

Next!

Spectral Plot Results Table Prediction Plots

Copy CSV Excel Search:

| Variable | Derivative | Pretreatment | Variable Selection | RMSEC | R2Cal | RPI |
|----------|------------|--------------|--------------------|-------|-------|-----|
| 1        | 2          | 2            | 1                  | 1.71  | 91.66 |     |

Showing 1 to 1 of 1 entries Previous 1 Next

"Prediction Plot" button is clicked to access obs-pred plots related to the pre-modeling result and graphics are displayed in the "Prediction Plots" sub-menu.

BAF

DATA INPUT

PRE-ANALYSIS

POST-ANALYSIS

ABOUT

**PRE-TREATMENT**

Derivatives: First (11 points)

Pretreatment: SNV (0 degree)

Spectral plot

**PRE-ANALYSIS**

Analysis 1: PLS (#comp.: 10, CV: 10%)

Apply Prediction plot

Next!

Spectral Plot Results Table Prediction Plots

Calibration Plot External Validation Plot

Model Prediction Reference Analysis

After the pre-analysis process is completed, you can switch to the "POST-ANALYSIS" menu with the "NEXT" button.

## Post-Analysis

POST-ANALYSIS menu is the screen where the predictive variable (wavelength) is selected based on the pre-analysis made in the previous menu and the final prediction model is created. The user writes the name of the target variable analyzed in the "VARIABLE NAME" section. The SelectWave implementation has four different options of filter methods and four options of wrapper methods for selecting the predictor variables. Along with no variable selection (None), total nine options are available under Variable Selection menu.

The filter methods are:

- VIP (Variable Importance on Projection)
- SR (Selectivity Ratio)
- sMC (significance Multivariate Correlation)
- mRMR (minimum Redundancy Maximum Relevance)

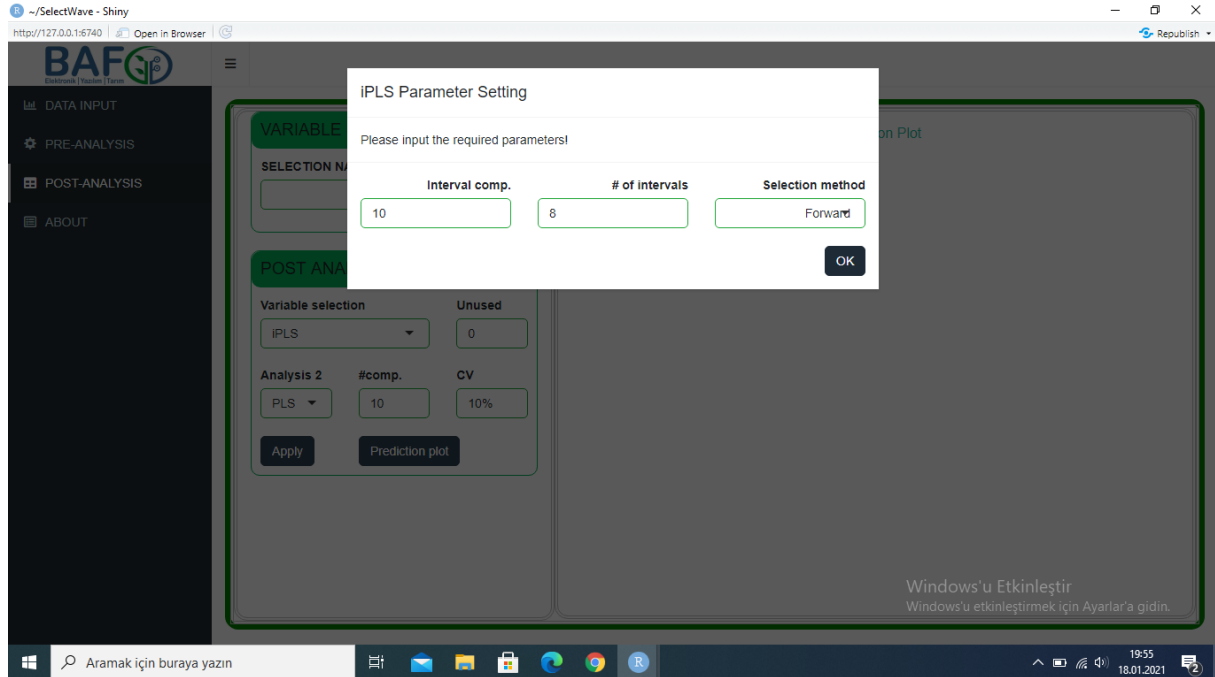
The wrapper methods are:

- iPLS (Interval Partial Least Squares)
- GA-PLS (Genetic Algorithm)
- IPW-PLS (Iterative Predictor Weighting)
- UVE-PLS (Uninformative Variable Elimination)

The selection is made from the "Variable selection" menu. Afterwards one assigns a "cut off" value for filter methods and input parameters for wrapper methods. Input parameters for wrapper methods are assigned in a pop-up window. There are different mandatory input parameters for each wrapper method.

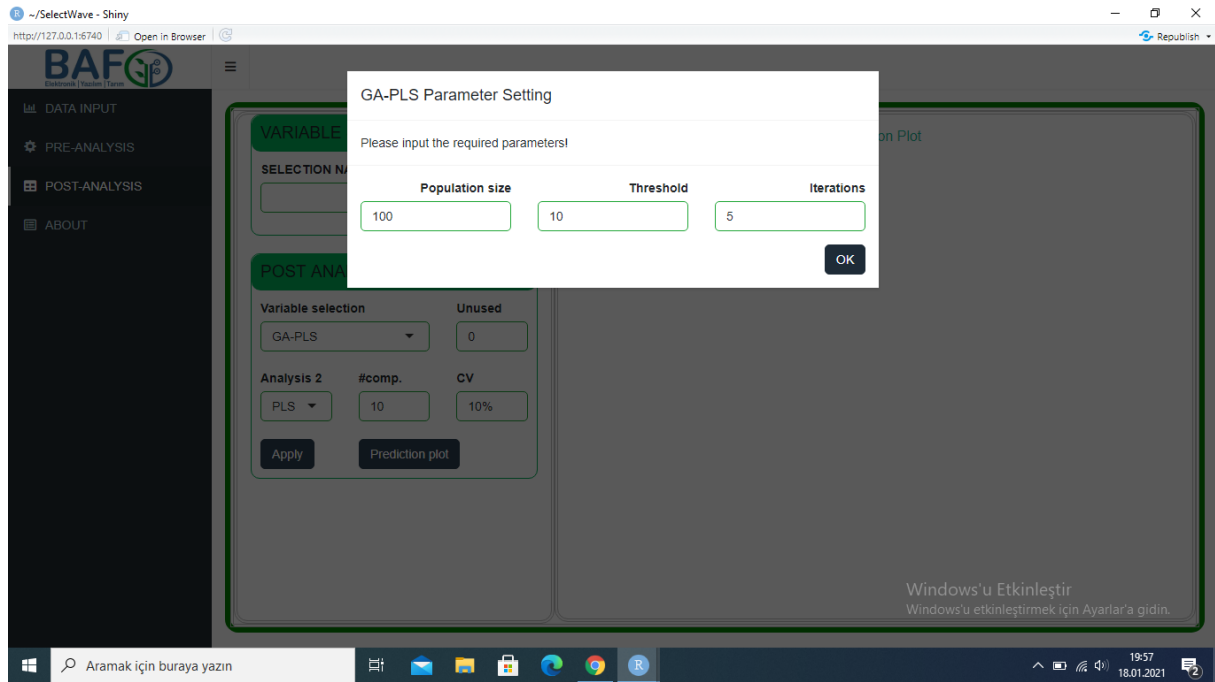
iPLS input parameters are;

- Interval components (Default value=10)
- Number of intervals (Default value=10)
- Selection method (Backward, Forward, Default value=Forward)



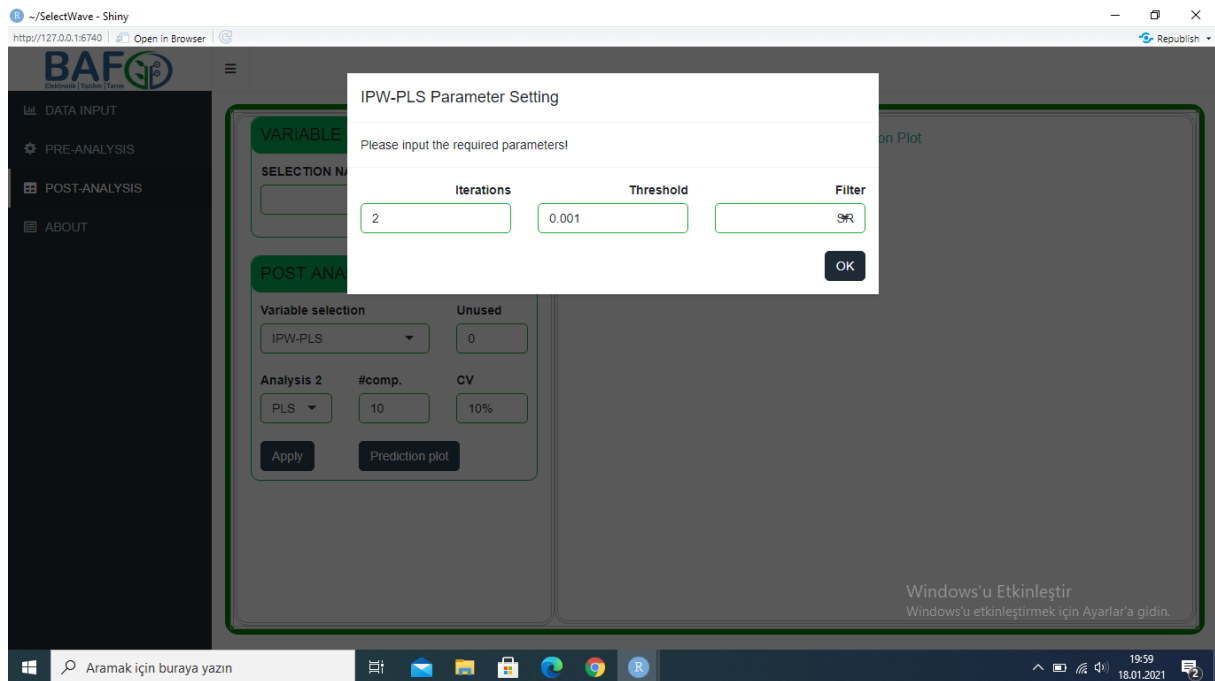
GA-PLS input parameters are;

- Population size (Default value=100)
- Threshold (Default value=10)
- Iterations (Default value=5)



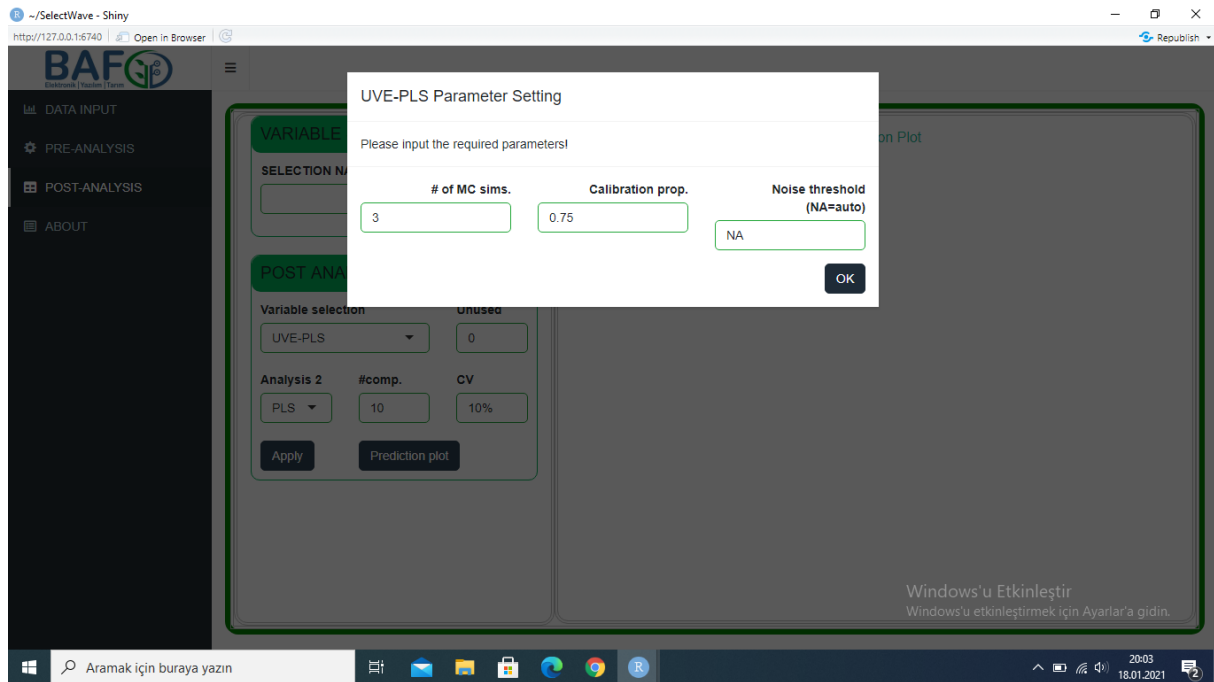
IPW-PLS input parameters are;

- Iterations (Default value=2)
- Threshold (Default value=0.001)
- Filter (Options=RC,SR,LW,VIP,sMC, Default value=SR)



UVE-PLS input parameters are;

- Number of MC simulations (Default value=3)
- Calibration prop. (Default value=0.75)
- Noise threshold (Default value=NA:Auto)

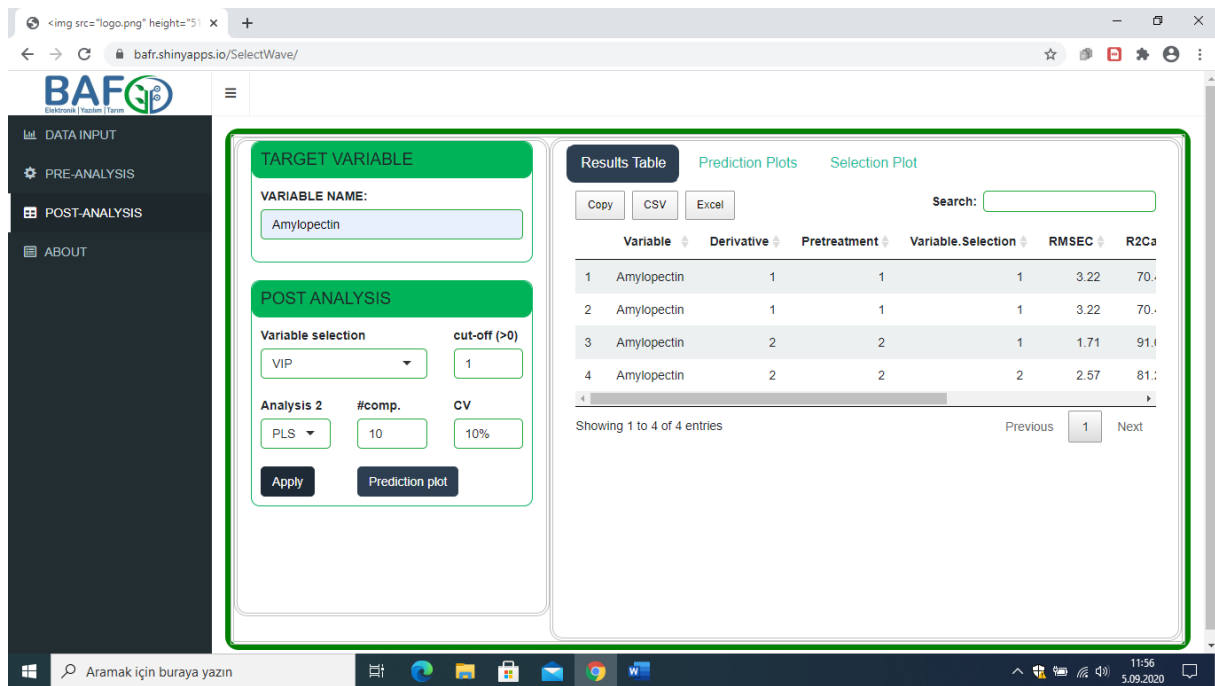


After assigning the all of the parameters required, the modelling method is selected. There are two options for the application as a modeling method.

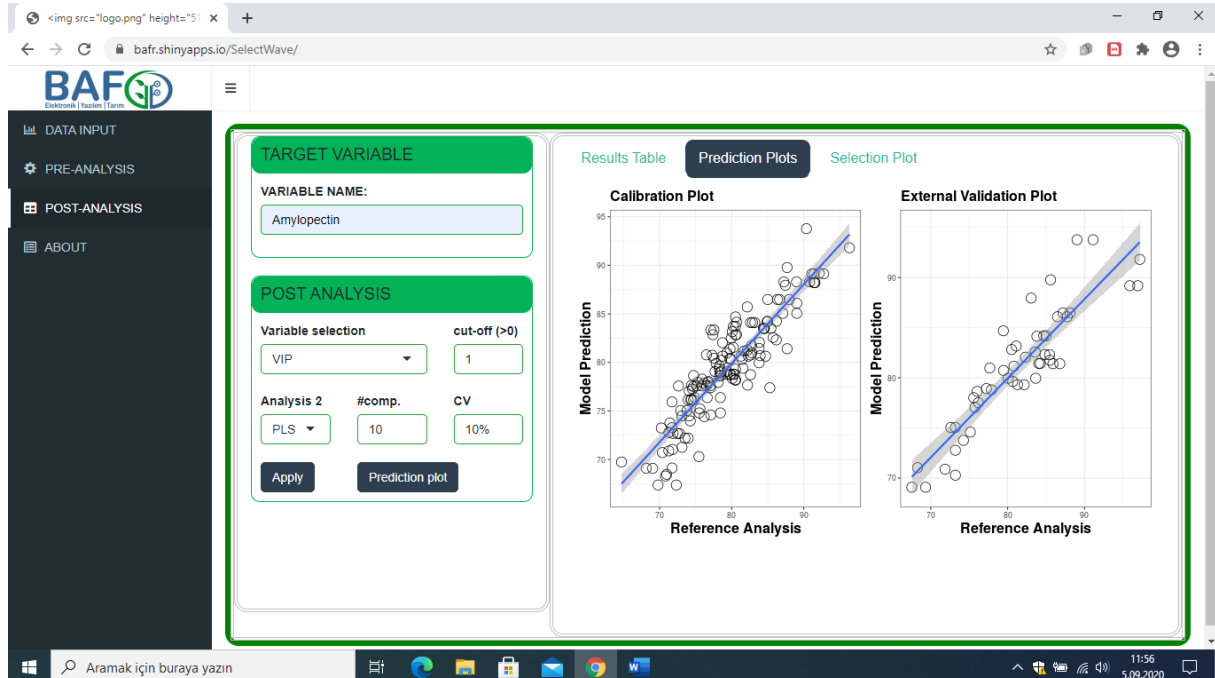
- PLS (Partial Least Squares Regression)
- SVM (Support Vector Machines)

After selecting the modeling technique and assigning the relevant parameters (COMP and CV for PLS, gamma and cost values for SVM), the analysis is performed by clicking the "Apply" button.

**Important:** If the PLS method will be applied in the analysis to be made in the POST-ANALYSIS menu in the SelectWave application, the COMP and CV values assigned in the PRE-ANALYSIS menu should be used.

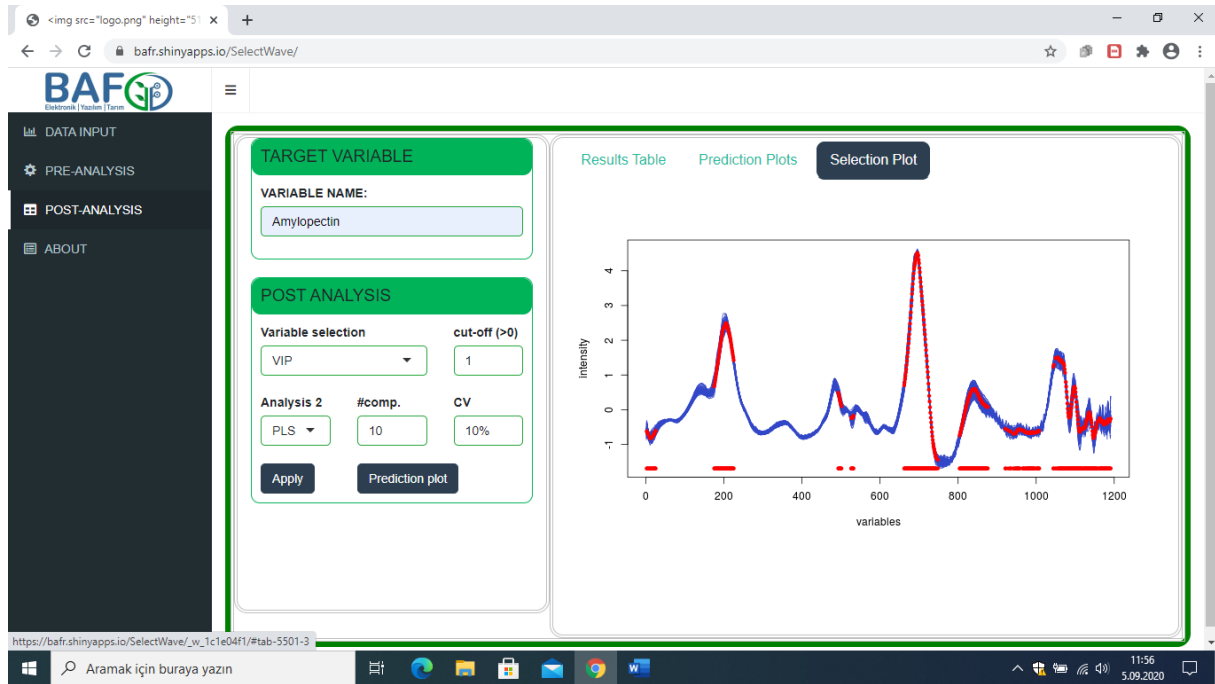


After the analysis results are reflected on the screen, the "Prediction Plot" button is pressed and the obs-pred graphics obtained from the calibration and external validation set of the model are displayed.



If the user has performed the analysis using a wavelength selection method, the application can automatically display the variables affecting the pre-processed spectral graph in red as a "Selection Plot" graph. User

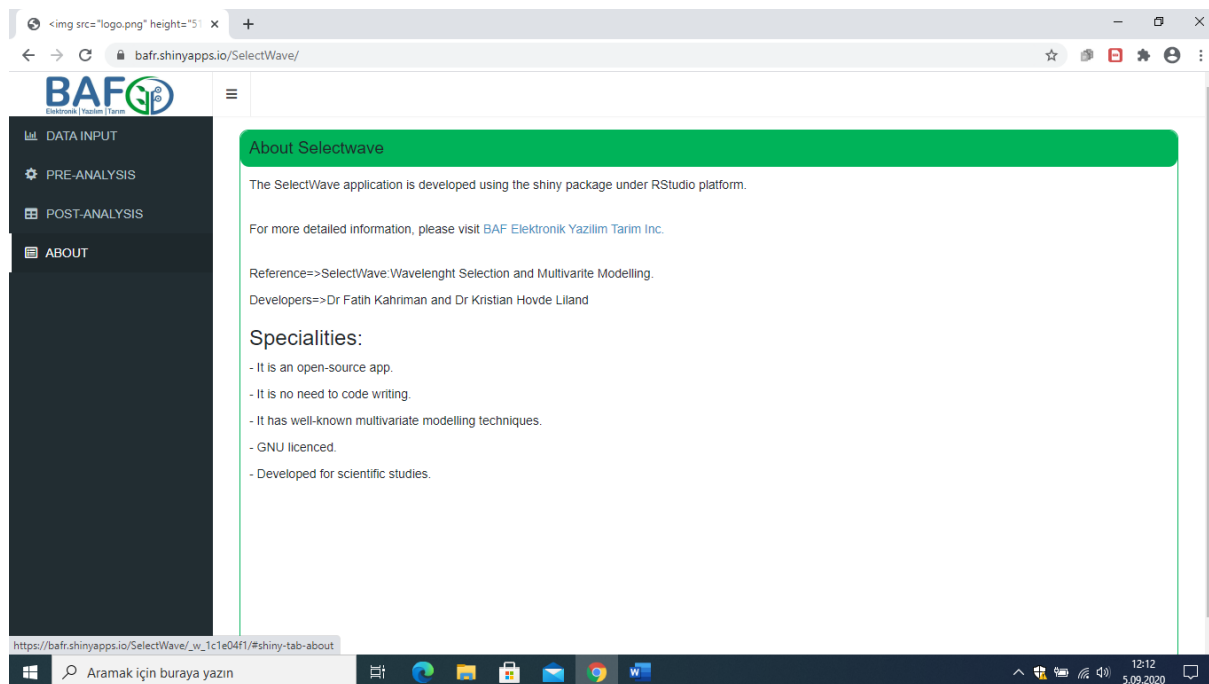
can view the results by clicking on the "Selection Plot" sub-menu after completing the analysis to view this graph. In the sample data used in the guide, after First Derivative + SNV transformation, VIP wavelength selection method and PLS modeling technique were applied as spectral pretreatment, the results are shown in the screenshot below.





## About

The last menu of the application contains the developers and general information about this application. This manual and some additional links can be accessed through this page.



For detailed information about the application, you can contact the developers at the following e-mail addresses.

Assoc. Prof. Dr. Fatih Kahrıman

e-mail: [fkahrıman@hotmail.com](mailto:fkahrıman@hotmail.com)

Assoc. Prof. Dr. Kristian Hovde Liland

e-mail: [kristian.liland@nmbu.no](mailto:kristian.liland@nmbu.no)

Citation Info: Kahrıman, F., Liland, K.H., 2021. SelectWave: A Graphical User Interface for Wavelength Selection and Spectral Data Analysis. Chemometrics and Intelligent Laboratory Systems, doi.xxxxxxxxxxxxxxxxxxxxxxxxxx. (<https://bafr.shinyapps.io/SelectWave/>)