

Software Description

MVC app: A smartphone application for performing chemometric methods



Hadi Parastar*, Hamidreza Shaye

Department of Chemistry, Sharif University of Technology, P.O. Box 11155-3516, Tehran, Iran

ARTICLE INFO

Article history:

Received 5 June 2015

Received in revised form 30 July 2015

Accepted 9 August 2015

Available online 15 August 2015

Keywords:

Multivariate calibration

Multiple linear regression

Principal component regression

Partial least squares

Smartphone

Android

ABSTRACT

In this work, a novel smartphone application entitled “MVC app” is developed to perform different multivariate calibration methods. This app is designed for chemists who are not expert in programming or in advanced statistics. The developed application can use any Android-powered device as an environment for running. It is an easy to use app which can simply install in your smartphone and play. Different multivariate calibration methods, such as multiple linear regression (MLR), principal component regression (PCR) and partial least squares (PLS) are included in this app. As an instance, for performing PLS modeling, first calibration and validation data sets are imported (via USB or Wi-Fi). Then, the number of latent variables (LVs) is chosen using leave-one-out cross-validation (LOO-CV). Afterwards, PLS model is built and the user can review the modeling results. In this regard, figures of merit (FOMs) of models, such as root-mean square error of prediction (RMSEP), standard error of prediction (SEP), bias and relative error (RE) and other parameters can be viewed for each analyte. Furthermore, various plotting options are included for each model. All of these options are available just by touching the screen, with no complexity that almost every chemist can use.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Calibration is an important topic in analytical chemistry which its purpose is determination of the amount of analyte(s) of interest in unknown samples in the presence of interferences. Moreover, as a side goal, it is sometimes convenient to obtain informative information about the identity of the analytes and the interferences, if present.

Owing to the fact that analytical chemistry can involve samples that are far from simple and often contain many components, therefore, multivariate calibration (MVC) methods have been proposed in recent decades to overcome fundamental mathematical challenges occurred during analysis of these complex mixtures [1]. The main goal of MVC is development of a quantitative model for prediction and interpretation of the properties of interest (e.g., concentration) using a number of predictor variables. Among different MVC algorithms, multiple linear regression (MLR) [2], principal component regression (PCR) [2] and partial least squares regression (PLSR) [3] has attracted great attention in chemistry in recent years due to their unique properties in the solutions and wide variety of applications. However, one of the most important features of chemometric methods is development of user-friendly software in order to develop the range of users of these methods. In this regard, development of software for chemists who are not expert

in programming or in advanced statistics will be an important mission for chemometricians. Great efforts have been done by different chemometric groups in this direction and different GUIs have been developed for performing MVC methods in user-friendly environments, such as PLS-toolbox [4], TOMCAT [5], MVC1 [6] and some other ones. On the other side, nowadays, smartphones are used as portable personal computers (PCs), changing our lives and how we interact to technology. The Android platform provides a vast range of applications that can be downloaded directly onto the phone. These mobile applications or “apps”, have a wide range of functionalities that can be used in many different disciplines [7]. They are easy-to-use, low price, flexible and finally anything is accessible just by touching the screen. Therefore, designing new applications for performing different types of chemometric methods can be considered as a new and interesting idea. In analytical chemistry some apps have been developed which are mostly used as transducers for colorimetric measurements by smartphone camera [8–10]. To the best of our knowledge, there is no Android-based application to perform chemometric methods. On the other side, these types of apps can be coupled to portable instruments which are highly useful for in situ and field studies. In these cases, data can be collected with instruments and then, data can be transmitted to a smartphone device through a Wi-Fi or USB connection and finally, they can be easily manipulated and processed with no extra effort.

Therefore, the aim of this work is development of a novel smartphone application (app) to perform different types of chemometric methods including multivariate calibration ones. The developed app

* Corresponding author. Tel.: +98 21 66165306; fax: +98 21 66029165.
E-mail addresses: h.parastar@sharif.edu, h.parastar@gmail.com (H. Parastar).

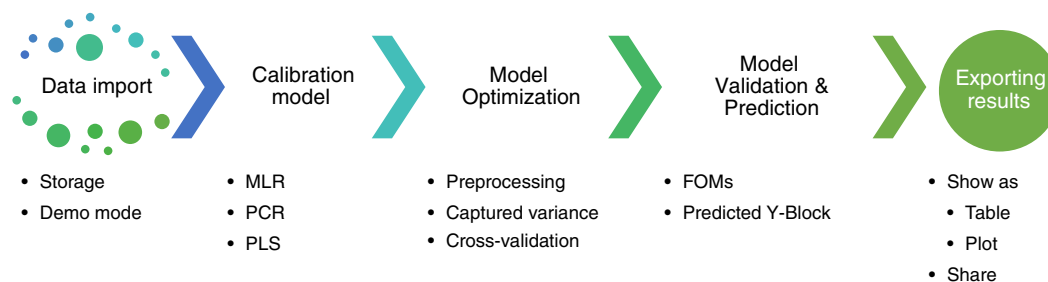


Fig. 1. General flowchart of MVC app for analyzing analytical data.

is based on Java language programming and it can be run on any hardware supported Java, here Android [11]. This app is called “MVC” (Multivariate Calibration) and it is designed for chemists who are not expert in programming or in advanced statistics and seek user-friendly tools for multivariate calibration. In MVC app, different MVC methods, such as MLR, PCR and PLSR are included. Also, several important tools for data preprocessing (e.g., mean-centring and auto-scaling), rank determination and model evaluation (e.g., Cross-Validation (CV) [12]) and plotting and sharing facilities are included. Running MVC app does not require a serious experience; however, a basic knowledge of the underlying methods is helpful to successfully interpret the results. More details about theoretical concepts behind the implemented methods can be found elsewhere [1–3,13].

With development of Smartphone apps, chemometric methods will be in your pocket and you can use them everywhere.

2. System requirements and installation

MVC requires an Android Smartphone with version 4.0 (IceCreamSandwich) or higher (the Android version can be checked from: Setting>About Phone>Android version). In order to import data and save logs as txt-file, MVC needs permission to read and write from phone external storage. This will be automatically reminded before installing MVC. MVC app can be installed by touching the app in your phone. MVC is available for free and you can access it by sending

email to the corresponding author of this paper (h.parastar@sharif.edu) and/or by going to our homepage [14]. In addition, there is an option to import data via MATLAB (Mathworks, Natick, MA, USA) [15] in which special function named “MVC.m” transform MATLAB’s matrices into .txt-file that can be simply copied into smartphone’s internal memory and imported by MVC [14].

3. Working with MVC app

Fig. 1 shows the flowchart of the MVC app to analyze different types of analytical data, such as spectroscopic, electrochemical and chromatographic ones. These steps are data import, selection of calibration model, optimization of calibration model, model validation and prediction of unknown samples and exportation of the results. These steps will be explained in the next section with a spectroscopic data set.

3.1. Preparing data

After launching MVC, the first page will be displayed as Fig. 2(a). Two options are available to continue. The first one is importing data from external storage by clicking “Load Data” button and the second one is to use implemented data by clicking “Demo” button. When using demo mode, the data will be imported automatically; containing calibration (25 samples), validation (7 samples) and test (5 samples) sets (see Tables S1–S4 of supporting information (SI)). The **X**-block

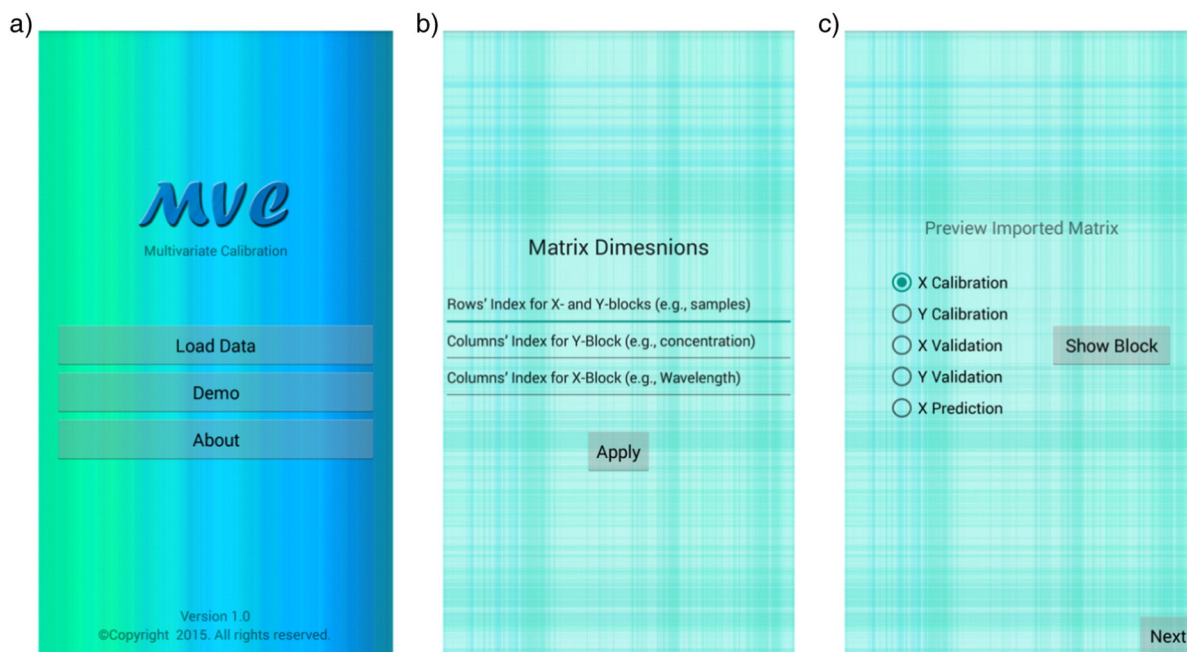


Fig. 2. (a) Starting page of MVC app after launching, (b) the MVC app page for dimensions indices and (c) preview imported matrix by MVC browsing tools.

for calibration set contains 25 samples that their UV–Vis absorption spectra were recorded in 27 wavelengths from 220 to 350 nm at 5 nm intervals. In addition, **Y**-block consists of 25 mixtures with different concentrations of 10 polycyclic aromatic hydrocarbons (PAHs) including pyrene (Py), acenaphthene (Ace), anthracene (Anth), acenaphthylene (Acy), chrysene (Chry), benzanthracene (Benz), fluoranthene (Fluora), fluorene (Fluore), naphthalene (Nap) and phenanthrene (Phen), respectively. Additionally, validation and prediction sets consist of first 7 and 5 samples of calibration set, respectively (Tables S3–S4 of SI). The corresponding data sets to test the MVC app were taken from ref. [2].

To load data from external storage, generated txt-file by “MVC.m” MATLAB-file can only be used. This function is as follows:

```
MVC('filename', XCal, YCal, XVal, YVal, XPred) (1)
```

where XCal and YCal are **X**- and **Y**-blocks for calibration set. Also, XVal and YVal are corresponding blocks for validation set and XPred as **X**-block of test set which is optional. After execution, a txt-file named “filename” made at MATLAB path. This txt-file can now be copied into smartphone external storage and then can be loaded into the app using “Load Data” option. After importing data, the indices for rows and columns of each data-block (i.e., **X** and **Y**) can be entered as Fig. 2(b). By double clicking on edit-text-boxes, the notation can be deleted. Note that if this step is skipped, MVC uses default dimensions indices.

Also, additional data are available by clicking “About” button in Fig. 2(a).

3.2. Browsing data

In each part of MVC, there are several tools to browse either imported or obtained data as a matrix. For example, when data imported into the MVC app, “preview imported matrix” page will be displayed as shown in Fig. 2(c). Note that “Preview imported matrix” page can be skipped by clicking “Next” button. All data blocks can be displayed after clicking “Show Block” button. In this regard, a new page named “Printing Matrix” will be displayed as depicted in Fig. 3.

In this new page, there are two options: (1) showing data as a table as demonstrated in Fig. 4(a) which it is possible to scroll display horizontally and/or vertically to scan all matrix data and (2) showing data

as a plot as depicted in Fig. 4(b). In this way, horizontal scrolling in bottom-one-third of screen will change across the rows of data matrix and tapping the screen by three fingers will share current plot as JPEG-picture. The path of the saved JPEG-file will display as a toast-message below screen.

It should be pointed out that the same procedure will be used all along the application to show data matrices.

3.3. Selecting and optimizing calibration model

After importing and reviewing data, calibration model can be built by using three popular MVC methods of MLR, PCR and PLSR as shown in Fig. 5(a). For the sake of brevity, here, only PLS modeling for UV–Vis data example will be discussed. As shown in Fig. 5(b), the **X**- and **Y**-block data sets can be preprocessed (mean-centring and autoscaling) before modeling [3].

One of the most important features in every calibration model is the number of significant latent variables (LV) for model construction which needs to be known before modeling. Among different methods, the captured variance by each LV [16] and cross-validation (e.g., leave-one-out) [12] are powerful techniques for this purpose which is included in MVC app. Therefore, after selecting appropriate preprocessing method, “Captured Variance” button should be clicked and then maximum number of PCs should be selected to show table of captured variance for each LV in **X**- and **Y**-block data sets in both individual and cumulative format as shown in Fig. 6(a). Also, “Cross Validation” button should be clicked to select the significant LVs according to the plot of RMSE-CV against number of LVs. Therefore, the maximum number of LVs to check should be entered. After performing CV for models with different number of LVs, the results are provided in both plot and matrix format. Fig. 6(b) shows the plot of RMSE-CV vs. number of LVs for one of the analytes (e.g., Nap).

It is noteworthy that when using CV in PCR method, the results are provided in terms of predicted residual error sum of squares cross-validation (PRESS-CV) vs. number of LVs for one of the analytes.

Finally, by clicking “Create Model” button, the desired model with specified preprocessing method and number of LVs will be created.

3.4. Validation and prediction, exporting and sharing the results

After model development, model validation and prediction of unknown samples using calibration model (e.g., PLSR) will be possible (Fig. 7(a) and (b)). Note that if there is no validation set in the imported data, MVC app will automatically select calibration set as validation set to calculate the statistical parameters. By clicking the “Figures of Merit” button in Fig. 7(a), FOMs including root-mean square error of prediction (RMSEP), standard error of prediction (SEP), bias and relative error (RE) are calculated for developed model (Fig. 8). In addition, the user can browse predicted **Y**-block as both plot and matrix formats in this step. Finally, the developed and validated calibration model can be used for prediction. In this manner, pass the validation step and go to the prediction page by clicking “Prediction” button.

Fig. 7(b) shows the prediction page of MVC app. As it can be seen, there are same features as previous step to browse predicted **Y**-block as plot and matrix. In addition, user can save logs as txt-file containing analysis method name, type of preprocessing for both data blocks, number of LVs, data for calibration, validation and prediction, predicted **Y**-block for validation and prediction steps and matrix of regression coefficients of MVC model (e.g., PLSR). Also, there is an option to share logs' file by smartphone via Bluetooth, message, Email and other supported methods depends on user's device. It is important to note that logs-file is just for users' notification and cannot be used as data input in import section. By clicking “Finish” button as in Fig. 7(b), current wizard will be closed by user confirmation.

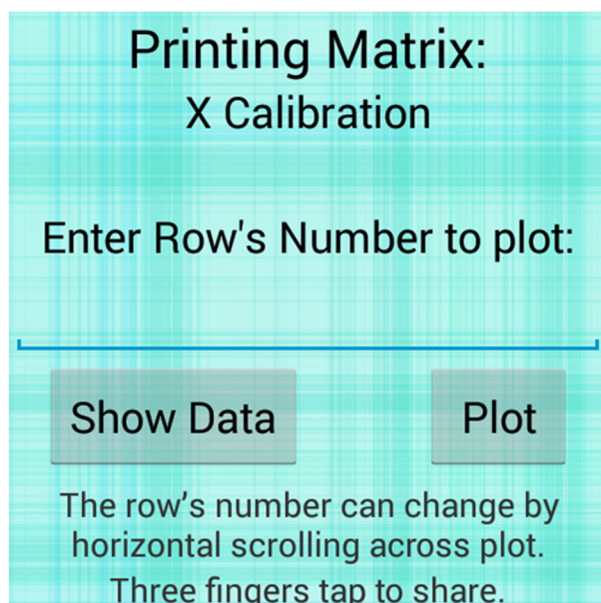


Fig. 3. Printing matrix page, choosing different types of plot browsing.

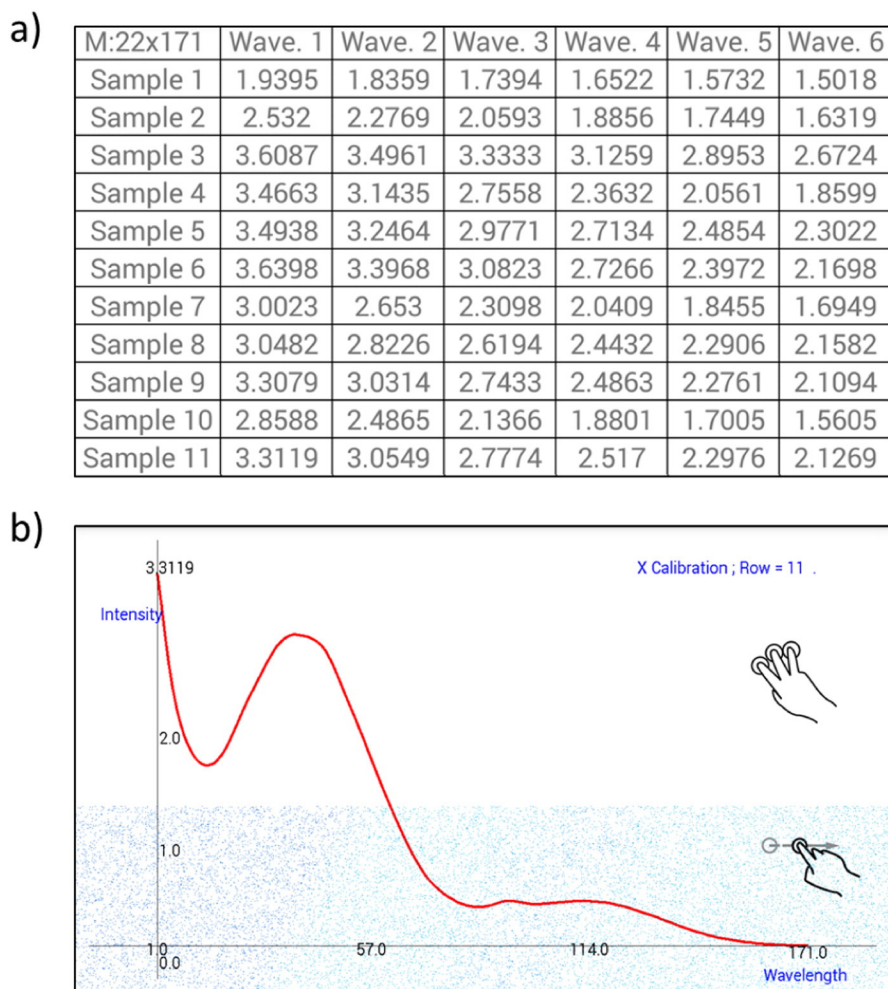


Fig. 4. Showing matrix (a) as a table, it is possible to scroll and move along rows and columns and (b) as a plot, it is possible to scroll horizontally to change rows in bottom-one-third of screen display and to tap by three fingers to share current plot as JPEG-file.

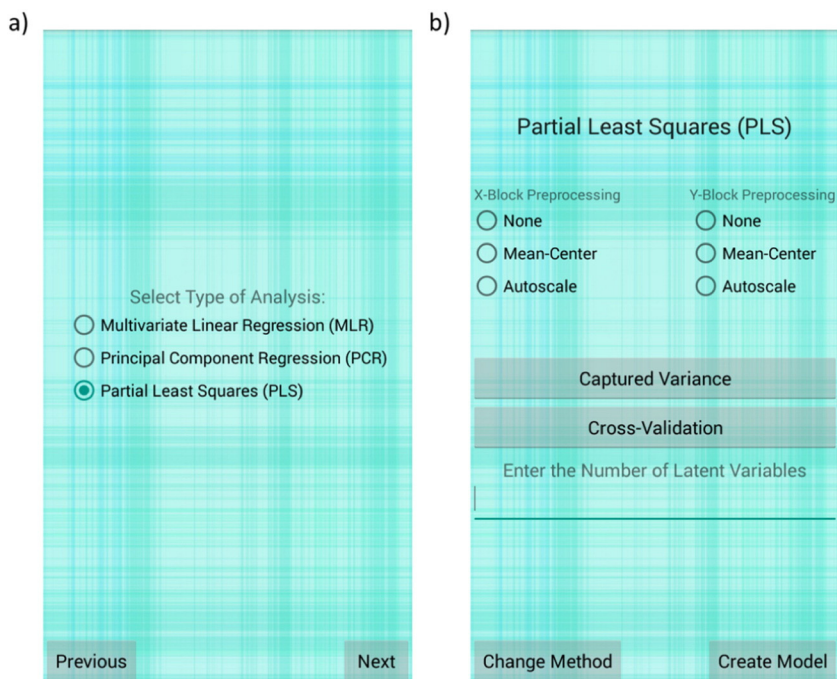


Fig. 5. (a) The page of selecting calibration model using MVC methods of MLR, PCR and PLSR. (b) The MVC app activity which provides several tools for PLSR, such as preprocessing and cross-validation and captured variance.

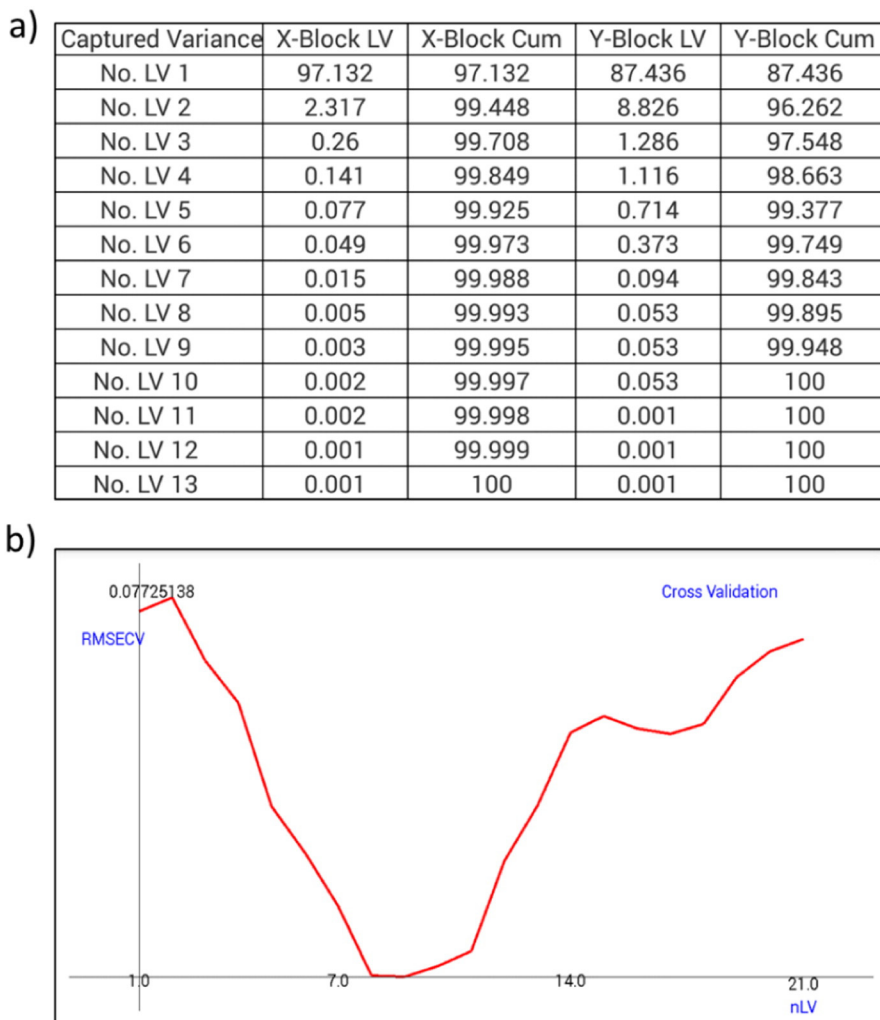


Fig. 6. (a) Table of captured variance by each LV for X- and Y-block data sets in both individual and cumulative formats. (b) Plot of RMSE-CV vs. number of LVs for Naphthalene to choose the significant number of LVs.

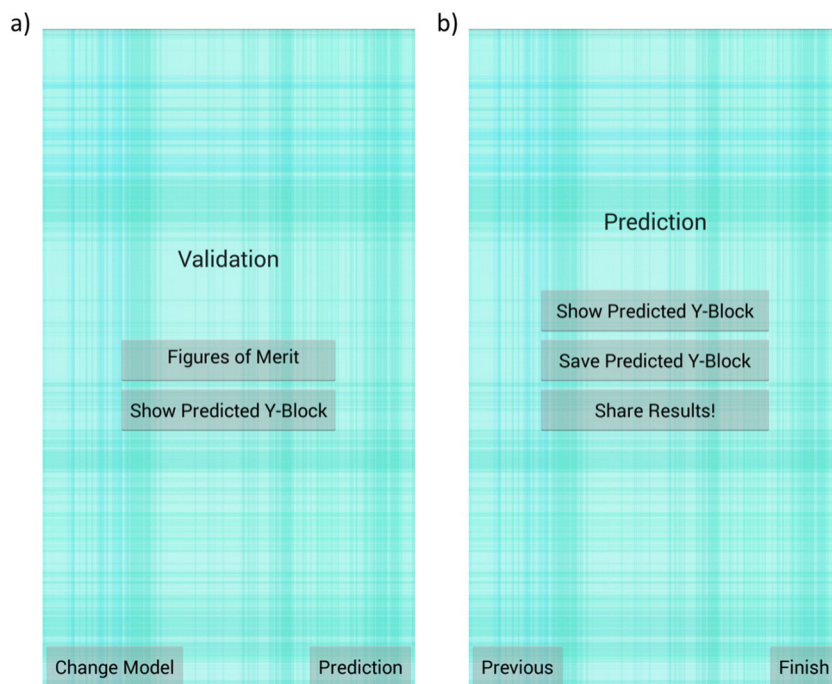


Fig. 7. (a) Validation and (b) prediction pages of MVC app.

FOMs	Num	RMSEP	SEP	Bias	Rel. Error (%)
Analyte:	1	0.031	0.033	-0.001	4.985
Analyte:	2	0.039	0.042	-0.001	24.237
Analyte:	3	0.02	0.021	-0.003	8.681
Analyte:	4	0.038	0.041	0.005	24.368
Analyte:	5	0.025	0.025	-0.009	5.933
Analyte:	6	0.052	0.054	-0.014	2.832
Analyte:	7	0.027	0.028	-0.008	20.832
Analyte:	8	0.104	0.112	0.004	16.185
Analyte:	9	0.019	0.021	-0.001	12.855
Analyte:	10	0.07	0.075	0.008	11.13

Fig. 8. Figures of merits (FOM) in demo mode, using autoscaling for both X and Y-blocks.

4. Conclusions

Nowadays, smartphones become a part of our lives and change how we look to life, science and technology. In this regard, many applications have been developed for different purposes which are related to information delivery. Indeed, these applications can play an important role in chemistry and especially in analytical chemistry. This would revolutionize portable analytical instrument usages. However, they should be as simple as possible that every non-expert user can use them. In this work, a novel Android application named MVC app was developed. The MVC app is a flexible and user-friendly application capable of performing multivariate calibration methods including MLR, PCR and PLSR. Also, performing different chemometric methods, such as pre-processing methods, CV, plotting and sharing results is possible in MVC app. It is important to notify that modern smartphones are fast enough to perform chemometric methods within a reasonable time scale (e.g., a few seconds) regardless of the data size. Owing to the fact that the usual matrix sizes for first-order analytical data are not too big, therefore, MVC app can handle them in a reasonable time period. However, for some special processing methods, such as cross-validation in both PLS and PCR methods, it takes a while to perform via smartphones which is reasonable due to its iterative nature.

It is concluded that development of user-friendly software and apps can significantly increase the rate of application of chemometric methods in different disciplines.

5. Independent testing

Prof. Alejandro C. Olivieri

Rosario Institute of Chemistry (IQUIR-CONICET), Suipacha 531, Rosario (2000), Argentina, e-mail: olivieri@iquir-conicet.gov.ar.

The authors have developed the Smartphone application "MVC App" which allows one to carry out various multivariate calibration methods. I have downloaded and installed the application on a cell phone and a tablet, both under Android as operating system, and found it to work as the authors describe. The developed application can implement multiple linear regression (MLR), principal component regression (PCR) and partial least-squares (PLS). The number of latent variables can be tuned using the well-known cross-validation. Saving data, prediction results and figures of merit reports can be easily obtained. I believe that this software may be highly useful for in situ and field studies, where data can be collected with instruments such as a near infrared spectrometer including a diode array detector. If data can be transmitted to a smartphone device through a wifi connection, they could then be easily manipulated and processed with no extra effort.

Prof. Mehdi Jalali-Heravi

Chemistry and Biochemistry Department, California State University, Los Angeles, Phone: (949) 466 4766, Email: mjalali2@calstatela.edu.

This software is a very interesting piece of work, especially in opening a new window for carrying multivariate calibration (MVC) methods using smartphones. I confirm all abilities of this smartphone application as authors described in the manuscript. This software helps those researchers who are interested to transfer a data to their smartphones through Wifi connection and manipulate and process them using MVC techniques. This APP is able to implement the common methods of multiple linear regression (MLR), principal component regression (PCR) and partial least squares (PLS). For each model, the statistical parameters, such as root-mean square of prediction (RMSEP), standard error of prediction (SEP) and relative error (RE) can be viewed. One of the advantages of this software is that it is very easy to be installed and its application is very simple and as mentioned in the manuscript no needs to be expert in programming. The presented smartphone MVC application works on different Android-powered devices. And I believe that regarding the easiness and simplicity of this App, all chemists who are interested in MVC analysis can use it.

Conflict of interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The authors would like to thank the Research Council of Sharif University of Technology (SUT) for their financial supports. Also, the authors would like to thank Ankit Srivastava for 2D plot class and EJML library developers (<http://ejml.org>).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.chemolab.2015.08.010>.

References

- [1] K.R. Beebe, B.R. Kowalski, An introduction to multivariate calibration and analysis, *Anal. Chem.* 59 (1987) 1007A–1017A.
- [2] R.G. Brereton, *Chemometrics: Data Analysis for the Laboratory and Chemical Plant*, John Wiley & Sons, 2003.
- [3] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, *Chemom. Intell. Lab. Syst.* 58 (2001) 109–130.
- [4] B.M. Wise, N. Gallagher, R. Bro, J. Shaver, W. Windig, R.S. Koch, *PLS Toolbox 4.0*, 3905, Eigenvector Research Incorporated, 2007.
- [5] M. Daszykowski, S. Serneels, K. Kaczmarek, P. Van Espen, C. Croux, B. Walczak, TOMCAT: a MATLAB toolbox for multivariate calibration techniques, *Chemom. Intell. Lab. Syst.* 85 (2007) 269–277.
- [6] A.C. Olivieri, H.C. Goicoechea, F.A. Iñón, MVC1: an integrated MatLab toolbox for first-order multivariate calibration, *Chemom. Intell. Lab. Syst.* 73 (2004) 189–197.
- [7] D. Libman, L. Huang, Chemistry on the go: review of chemistry apps on smartphones, *J. Chem. Edu.* 90 (2013) 320–325.
- [8] H. Yu, Y. Tan, B.T. Cunningham, Smartphone fluorescence spectroscopy, *Anal. Chem.* 86 (2014) 8805–8813.
- [9] B.-Y. Chang, Smartphone-based chemistry instrumentation: digitization of colorimetric measurements, *Bull. Korean Chem. Soc.* 33 (2012) 549.
- [10] F.M. Walker, K.M. Ahmad, M. Eisenstein, H.T. Soh, Transformation of personal computers and mobile phones into genetic diagnostic systems, *Anal. Chem.* 86 (2014) 9236–9241.
- [11] Google Mobile, <http://www.google.com/mobile/android/>.
- [12] R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection, *Int. Joint Conf. Artif. Intell.* 14 (1995) 1137–1145.
- [13] T. Næs, H. Martens, Principal component regression in NIR analysis: viewpoints, background details and selection of components, *J. Chemom.* 2 (1988) 155–167.
- [14] MVC app version 1.0, <http://sharif.edu/~h.parastar/software.html>.
- [15] MATLAB, <http://www.mathworks.com/>.
- [16] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, *Chemom. Intell. Lab. Syst.* 2 (1987) 37–52.