**Calif - the interactive Shiny web application for calibration of weights of statistical surveys**

Boris Frankovič, Statistical Office of the Slovak Republic, boris.frankovic@statistics.sk

**Abstract**

*Since many years the Statistical Office of the Slovak Republic has performed calibration of weights in statistical surveys in order to enhance precision requirements and maintain consistency between surveys. Based on our best practice, it has been proved that this procedure has to be carried out with utmost care. To that end, we prepared the first version of Calif with graphical user interface in 2013. This tool enabled users to take advantage of simple environment full of useful statistics that allowed for swift calibration when the most feasible solution was easily found. However, the old-fashioned GUI has found its limits. Therefore, we are coming with the new interactive web application made in R package Shiny that offers space not only for the enhancement of the calibration process but also for further development. Apart from new graphical user interface that runs in a web browser, makes calibration an intuitive and convenient process and offers new features, Calif could be used in the future as a more general statistical tool, which would not just be limited by the calibration but also be open for imputation, estimation and variance calculation.*

**Keywords:** calibration, weights, Shiny, R

**1. Calibration approach**

Calibration of weights in social and business surveys is some of the main topics of the Methodology Department of the Statistical Office of the Slovak Republic. This procedure is able to significantly enhance precision of survey estimates, especially in cases when study variables are highly correlated with calibration auxiliary variables and response rates among certain population groups, which are covered by calibration variables, are unequally distributed. Moreover, calibration introduces desired consistence among several surveys. Without calibration, parameters derived from statistical surveys are just the estimates of population characteristics, with some standard errors. However, if these population values are known at some level, they can be used to adjust the survey weights so that the population estimates comply with their real values. As Sarndal (2007) states, “Calibration is a procedure that can be used to incorporate auxiliary data. This procedure adjusts the sampling weights by multipliers known as calibration factors, that make the estimates agree with known totals. The resulting weights are called calibration weights. These calibration weights will generally result in estimates that are design consistent, and that have a smaller variance than the Horvitz-Thompson estimator.”

According to Calif manual (2018), the procedure can be described as follows:

Let us consider a population with units, from which the probability sampling of size is undertaken. Every unit in has design sampling weight, which is equal to where is the inclusion probability of unit , possibly adjusted for nonresponse. The objective is to estimate the population total of a study variable , denoted as by means of the Horvitz-Thompson unbiased estimator . Assume auxiliary variables and their population totals . These can be known from administrative sources, censuses or even broader surveys. The main objective of the calibration approach is to reproduce the new weights for each that confirm auxiliary totals and differ minimally from design weights . Then, these new weights, called calibration weights, can be used to estimate any parameter from any study variable in the survey.

Let denote the calibration weight of element . The calibration estimator of a study total is

|  |  |
| --- | --- |
|  | (1) |

while calibration constraints are fulfilled

|  |  |
| --- | --- |
|  | (2) |

for all .

The distance between design and calibration weights is expressed via so-called distance function. Let denote the quotients of these weights (known as calibration factors or g-weights). Then the distance function is a nonnegative convex function of with minimum in 1 (so when calibration and initial weights are equal). As stated in Frankovic (2013), to find calibration weights we have to find a minimum of the equation

|  |  |
| --- | --- |
|  | (3) |

where , , , is a vector of Lagrange multipliers and is a matrix of auxiliary variables.

By taking partial derivatives of we get

|  |  |
| --- | --- |
|  | (4) |

where is the inverse function to derivative of . This system can be solved by several optimization methods taking as starting values.

The choice of a proper distance function constitutes a crucial step in calibration process. From among the large set of functions that fulfill the above-mentioned criteria, four are commonly used. They are:

1. *linear* - the basic function, if some solution is algebraically possible, with the use of this function it will always be found. On the contrary, it can introduce negative weights, which is infeasible in statistical production. It is defined as

|  |  |
| --- | --- |
|  | (5) |

1. *raking ratio* - exponential function that circumvents the negative weights aspect, however giving weights between 0 and 1, which is also inappropriate for state statistics. Sometimes the average difference between initial and calibration weights is lower than with the linear function. Its definition is

|  |  |
| --- | --- |
|  | (6) |

1. *linear bounded* - bounded version of linear function, when user is able to enter the interval for calibration quotients (g-weights, as defined above) and thus take control of the process. This is very useful in practice, as user can obtain the new weights within some bounds, if possible, avoiding implausible weights. It will be , where and are the lower and upper bounds, respectively. User must be aware of range allowed for calibration weights, tense bounds often lead to unsolvable system and increase average difference applied to each initial weight .
2. *logit* - bounded version of raking ratio. In contrast to linear bounded method, the domain and range of the function directly contains the bounds.

|  |  |
| --- | --- |
|  | (7) |

**2. Calif**

Several tools deal with calibration. One can find many more or less usable programmes that differ in many aspects. Some of them run in SAS, however, we had put our focus on those that are free to use, thus started to use some calibration functions in R software. In order to simplify the work and make it effective, comfortable and powerful, we made the first version of Calif back in 2013. It offered user-friendly GUI environment, was free to use and very powerful in seeking appropriate and even approximate solutions (some tools can find just the accurate solution but it rarely exists, especially for many auxiliary constraints). Over time, some new versions arrived, however, with no major development, so they reached their limits in graphical user interface appearence and operability, which could have discouraged some users to work with it. Despite the functionality of the last version, Calif 3.3, some notable innovation was sought.

Calif 4.0 is a new Shiny web application with modern and attractive design. It is very easy to use and very fast, offers many features that can help users to find the best solution whilst maintaining time-proven techniques. The whole application is built in R software under the **shiny** package, while incorporating *calib* function from package **sampling** together with function *nleqslv* from package **nleqslv**. The diversity of ways how to find a good solution makes Calif a very interesting and comfortable tool. The various options of Calif require from the user some level of expertise. However, the easy-to-use graphical user interface makes it intuitive and comfortable to work with. Calif runs in several web browsers locally, without any concerns of leaving sensitive data outside currently used PC. No R knowledge is necessary.

It can be downloaded or run directly from GitHub repository <https://github.com/SO-SR/Calif> or from the SO SR’s webpage. Besides the source code, on the GitHub page you can find all necessary information including the Manual, running instructions, link to the SO SR’s webpage, screenshots and demo data.

The main features of Calif 4.0 are:

* web-designed graphical user interface
* two-stage calibration
* interactive data loading
* stratification
* detailed and user-defined specification of calibration variables
* powerful and fast solvers
* approximate solutions
* commonly used distance functions
* bookmarking
* graphical and worthful outputs
* and many others

In order to run Calif properly, the newest version of your web browser is needed. After running Calif from R (with one command; see GitHub page) you will be redirected into your web browser that will act as a display device, while all computations as well as loaded data remain locally in R; no data are sent to the Internet. Calif workspace is divided into three parts, the Overview tab, the Data tab and the Calibration tab.

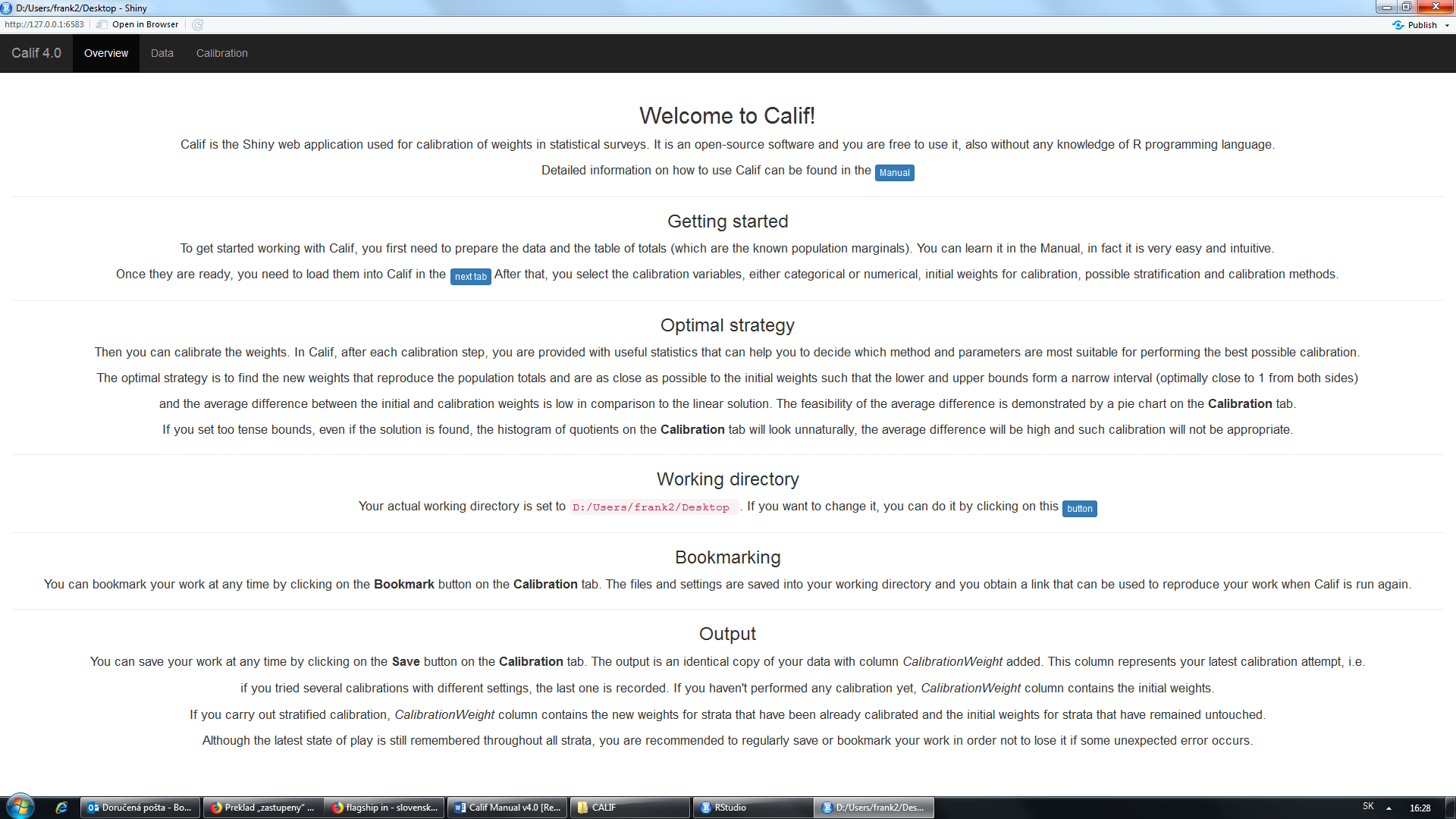
*2.1. Overview tab*

The first thing you can see after running Calif is the Overview tab. It displays main information on Calif, optimal calibration strategy, set up of the working directory, where the output files will be saved and some comments on bookmarking and output format. This tab can provide you with the general know-how you need to refresh from time to time, without reading the Manual.

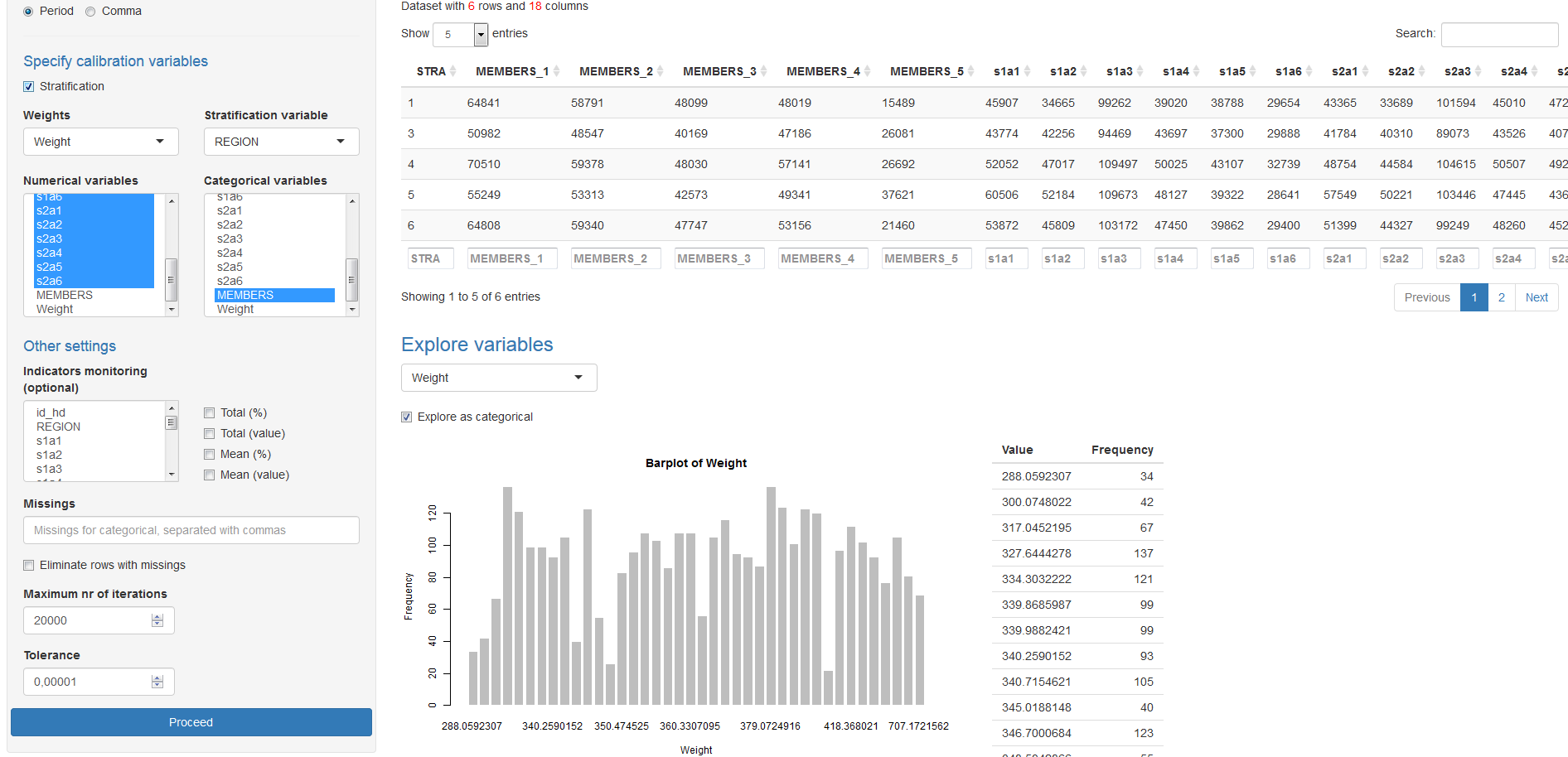
*2.2. Data tab*

On the Data tab, you can import the data and the table of auxiliary calibration totals and specify the calibration variables. Allowed formats are .csv, .txt or .sas7bdat. Data are interactively displayed in the main window, so you are able to visually check correctness of the import. You can switch between the single-stage and two-stage calibration, depending on the nature of your data (two-stage typically for household + individuals data).

**Figure 1. The Overview tab**



**Figure 2. The Data tab**



Required structure of the data is quite free for Calif, however, the table of totals has to follow a pre-defined layout. Although this is not the subject of this paper, you can examine the requirements in the Manual.

It is necessary to tell Calif which variables are deemed auxiliary numerical or categorical, which designate strata allocation, initial weights or possibly household ID. All of these can be set on this tab. Due to this utility, imported data can contain many variables that are not relevant for calibration; they will be just left out of the process but you are not obliged to delete them exclusively for Calif import. Moreover, you are free to explore any variable of the data, either as categorical or as numerical, simply by using the tool on this tab.

*2.3. Calibration tab*

This is the main part of Calif, where you can choose proper method, solver, calibrate the weights, see the outputs and decide, which solution is most feasible. If you have set stratified calibration prior to this tab, the strata list appears. There you can select various strata and calibrate them separately to each other. Just bear in mind that Calif always remembers the last calibration along with its settings that has been performed in each stratum. If you omit some stratum, its weights remain unchanged. If you calibrate some stratum several times, only its last setting is remembered. Further calibration of another stratum will not affect the previous result obtained for different stratum.

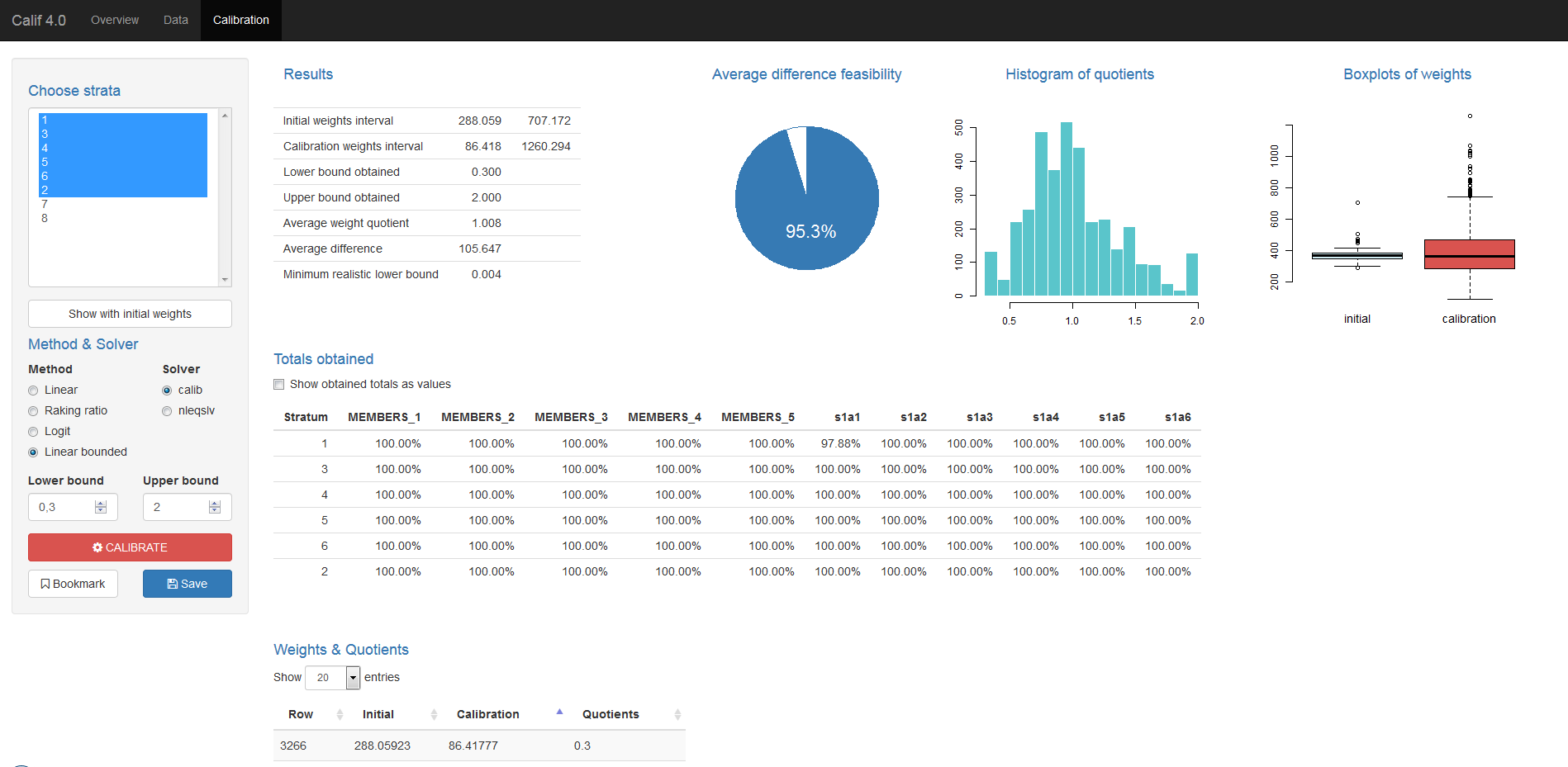
In statistical surveys, it often happens that the survey data are affected by unequally distributed non-response. Then, some groups are represented by more sample units than the others. Although this can be resolved by non-response analysis and weight adjustment, some differences still remain. In order to see the Horvitz-Thompson estimates of population calibration totals (prior to calibration), you can click the *Show with initial weights* button and check possible incorrectness in the data. Displayed totals should be ideally around 100%.

You have to select one method (distance function) and a solver. Default settings are the linear method and calib solver, but this combination often yields negative weights. Therefore the bounded methods are more appropriate. The bounds are expressed by proportion “calibration weight/initial weight“. The choice of proper method will be the topic of the next chapter.

Two powerful optimizers are available in Calif. Function *calib* (from package **sampling**) is very fast and powerful, function *nleqslv* (package **nleqslv**) is a little bit slower, however in some scenarios can perform better – especially in business surveys with numerical variables, when small strata with just a few units are calibrated and differences between H-T estimates and known population totals are significant. For social surveys (strata with many units and auxiliary variables) calib performs better. Each of the solvers is able to find also an approximate solution (not only exact).

After clicking the *Calibrate* button, resulting totals are shown, along with other statistics, including some plots, weight intervals, bounds obtained, average difference and minimum realistic lower bound. So displayed is a table with original and new weights. All of these outputs can assist you in finding the optimal solution.

**Figure 3. The Calibration tab**



You can save you work at any time during the process. Saved are the original data enriched with one additional column and the calibration settings. In stratified case, the resulting files will contain the calibration performed in each stratum, whereas the untouched strata will have the initial weights in the additional column. In contrast to saving the outputs, you can simply click the *Bookmark* button and obtain an URL, which can reproduce the current state during the next session.

**3. Optimal strategy**

The choice of a proper method and application of a sound technique is necessary when seeking the optimal trade-off as different settings may yield different results. The following is our best practice underpinned by theoretical background.

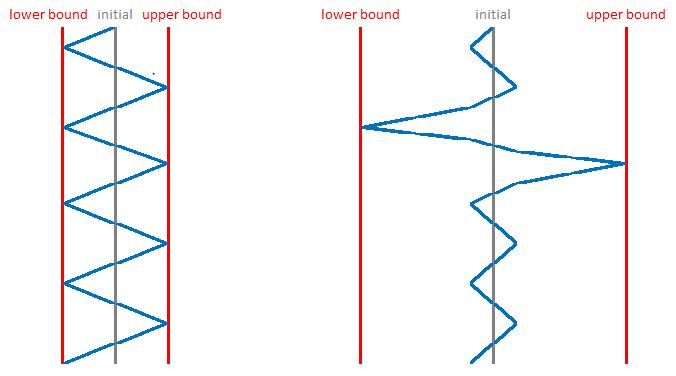
The optimal strategy is to find the new weights that reproduce the population totals and are as close as possible to the initial weights such that the lower and upper bounds form a narrow interval (optimally close to 1 from both sides) and the average difference between the initial and calibration weights is low in comparison to the linear solution. In order to fulfill this strategy, you should follow the next steps:

1. Linear method is run. Despite negative weights, this calibration is optimal, yielding minimal average difference. If resulting weights are within some reasonable interval, you can accept this solution. If not, continue to step 2.
2. Raking ratio method is run. If still not acceptable, continue with step 3.
3. Bounded method is run. You should try several bounds such that the population totals are reproduced, the bounds interval is as narrow as possible whereas the average difference is low compared to the linear method and weight quotients should not be pushed off to the bounds.

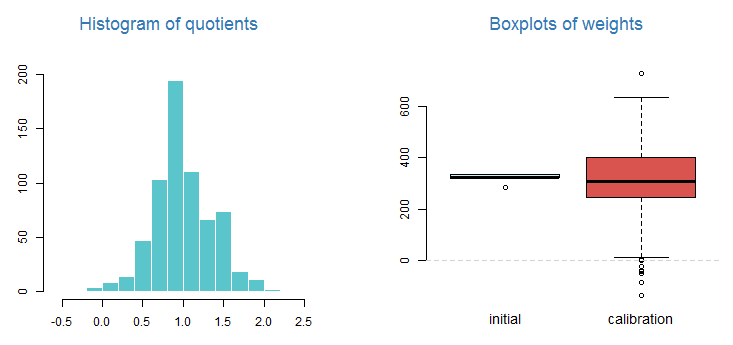
A common rule has been to find such solution that bounds are as strict as possible. In fact, this is not particularly true. Constrained bounds can keep calibration weights within some tense interval and thus creating a perception of a tiny difference between initial and new weights. However, limited space forces the initial weights towards bounds and as a result the average weight difference is quite high, i.e. the applied change is all but not tiny. On the other hand, less strict constraints form some area to move, where just a small number of weights is pushed off to the bounds and others remain close to their origins. Most obvious case is the use of linear method - no bounds presenting an open field cause a big difference for a few weights and a negligible difference for a vast majority of them - their distribution resembles the Gaussian curve. For that reason, the average difference between initial and calibration weights is minimal for unbounded methods (linear and raking ratio). Owing to the ambition to avoid extreme weights, we should impose some bounded method, but a feasible compromise between the width of the weight interval and the distortion applied has to be sought after.

To this end, there are several tailored outputs in Calif. First of all, an *Average difference feasibility* pie chart is included that calculates the ratio of average differences between linear and current solution. It should be as close as possible to 100%. Secondly, a *Histogram of quotients* and *Boxplots of weights* illustrate the distributions of weight quotients, initial and calibration weights. They can deliver a very nice outlook on the current situation - whether the weight quotients follow more or less normal distribution and new weights do not substantially differ from initial weights.

**Figure 4. Distortion of initial weights with strict vs. relaxed bounds**



**Figure 5. Histogram of quotients and Boxplot of weights**



**4. Conclusion**

In this paper we have introduced Calif - a new web-designed tool for calibration of weights of statistical surveys. It is completely made in R software and free to use. Over several years of experience in this field we have collected all the best practices and methodological findings to put them into this tool. User-friendliness, various utilities, comprehensive documentation and great operability make Calif a powerful calibration tool, which is well-suited also for users with lack of programming skills. The structure of Calif enables to extend it to a more general statistical tool, which would not be limited by just the calibration but also be open for imputation, estimation and variance calculation.

**5. References**

Deville, J.-C., Sarndal, C.-E. 1992. Calibration estimators in survey sampling. Journal of the American Statistical Association, 87, 376-382.

Sarndal, C.-E. 2007. The calibration approach in survey theory and practice. Statistics Canada, Business Survey Methods Division. Catalogue no. 12-001-X, Vol. 33, No. 2, pp. 99-119.

Harms, T., Duchense, P. 2006. On calibration estimation for quantiles. Survey Methodology, 32, 37-52.

Frankovic, B. 2013. Calibration of weights of statistical surveys in R language. Bratislava: Forum Statisticum Slovacum 5/2013, p. 19-37.

Sautory, O. 1993. La macro CALMAR. Paris: INSEE. Available at: <http://www.restore.ac.uk/PEAS/ex1datafiles/data/doccalmar.pdf>

Sautory, O. 2003. A new version of the Calmar calibration adjustment program. Statistics Canada International Symposium Series – Proceedings.

Vlacuha, R., Frankovic, B. 2015. The Calibration of Weights by Calif tool in the Practice of the Statistical Office of the Slovak republic. Bucharest: Romanian Statistical Review 2/2015, The [International Conference New Challenges for Statistical Software - The Use of R in Official Statistics](http://www.revistadestatistica.ro/index.php/the-international-conference-new-challenges-for-statistical-software-the-use-of-r-in-official-statistics/), paper.

Frankovic, B. 2018. Calif Manual v4.0. Statistical Office of the Slovak Republic. Available at: <https://slovak.statistics.sk/wps/wcm/connect/8a164814-ebda-43c1-b88e-28d40cb63a77/Calif+Manual+v4.0.pdf?MOD=AJPERES&CVID=m80Jevi&CVID=m6Tfvcg&CVID=m6Tfvcg&CVID=m6Tfvcg&CVID=m6IKz0d&CVID=m6IKz0d>

R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>

Winston Chang, Joe Cheng, JJ Allaire, Yihui Xie and Jonathan McPherson. 2017. shiny: Web Application Framework for R. R package version 1.0.5. <https://CRAN.R-project.org/package=shiny>

Berend Hasselman. 2014. nleqslv: Solve systems of non linear equations. R package version 2.1.1. <http://CRAN.R-project.org/package=nleqslv>

Yves Tillé and Alina Matei. 2013. sampling: Survey Sampling. R package version 2.6. <http://CRAN.R-project.org/package=sampling>

Hadley Wickham and Evan Miller. 2015. haven: Import SPSS, Stata and SAS Files. R package version 0.2.0. <http://CRAN.R-project.org/package=haven>