faoswsStandardization:
`Full Standardization and Balancing`
Data-sets content and plug-in execution

Cristina Muschitiello
Food and Agriculture Organization of the United Nations

21 June 2018

Abstract

This vignette provides a description on the executin of the "Full Standardization and Balancing" plugin: this is the plugin that, starting from the data collected and pulled into the input data-set `sua_unbalanced`, performs all the steps of the standardization and balancing (as described methodologically in a separate document) and save the data into 3 different output data-sets

Contents

List of Figures

Disclaimer

This Working Paper should not be reported as representing the official view of the FAO. The views expressed in this Working Paper are those of the author and do not necessarily represent those of the FAO or FAO policy. Working Papers describe research in progress by the authors and are published to elicit comments and to further discussion.

This paper is dynamically generated on June 21, 2018 and is subject to changes and updates.

The Data flow

The data Flow of the Standardization and Balancing Plugin is reported in 1. For detail about the methodology, please see the document *Standardization & Balancing for Food Balance Sheet Calculation*.

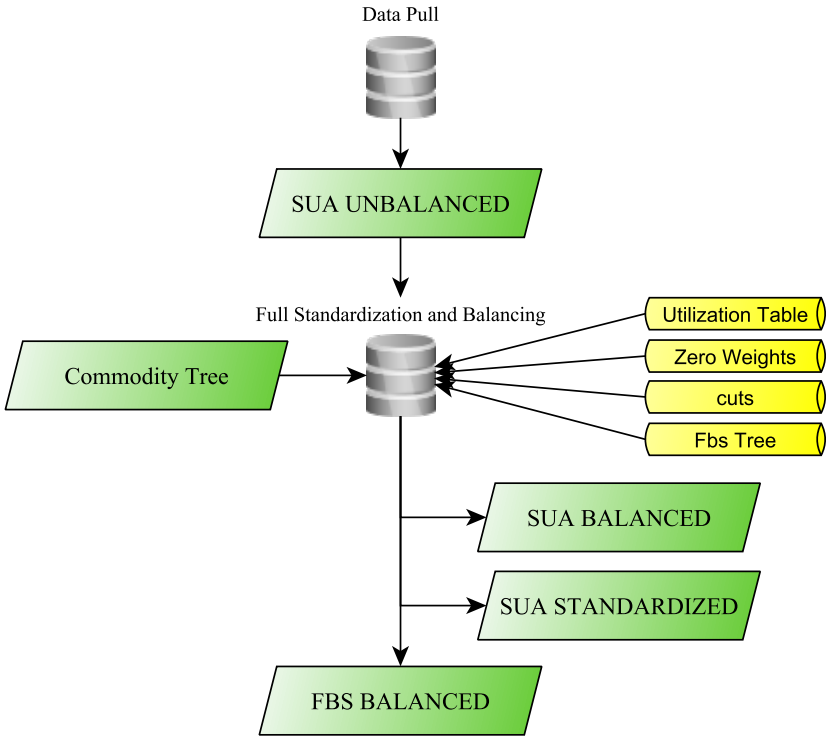


Figure 1: Data Flow of Standardizatrion and Balancing

The *Standardization and Balancing* involves 5 *datasets* and 4 *data tables* in the SWS. One peculiarity of this plug-in is that is saves data in 3 different data-sets. As a consequence, for executing it, it is necessary to open 3+1 sessions (3 for the output data-sets and 1 for the input data-set).

1 Log-in in the SWS

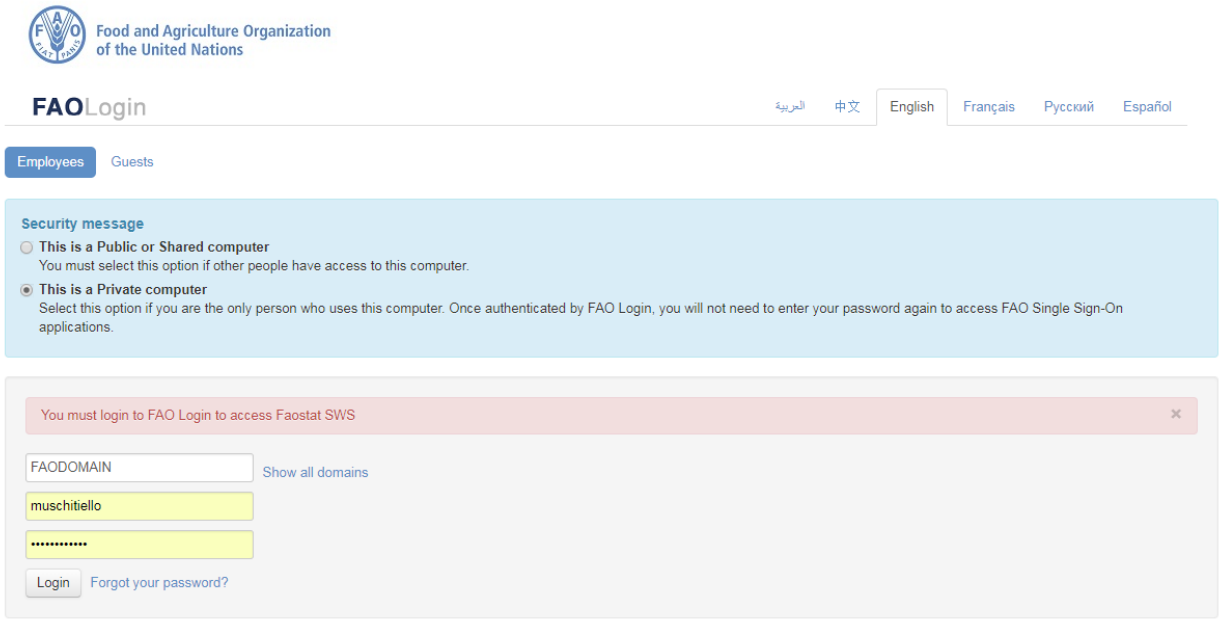


Figure 2: Log-in in the SWS

2 Data Pull

First. data from different data-sets have to be pulled inside the `sua-unbalanced` data-set and save the data back. This step is performed trough a plug-in called `pullDataToSua` which is documented in a separate document. A general workflow would probably start from the pulling pf all data for all countries, which are then saved in the SWS for all the users to start producing FBSs on single countries.

3 Open The Sessions.

4 sessions have to be opened, each one has to be named. This is not mandatory, but is important enough for reducing confusion and risk error when the plug-in has to be run. For this document an example on China Mainland, years from 2010 to 2016 is used.

3.1 `sua-unbalanced` session

This is the session on the *input* data-set. After having used the *New-session* button, this session has to be opened in the *suafbs:sua_unbalanced*. Therefore *SUA/FBS* domain and *sua_unbalanced* have to be selected from the screen (figures 3 to 5).

Figure 3: Open a new Session

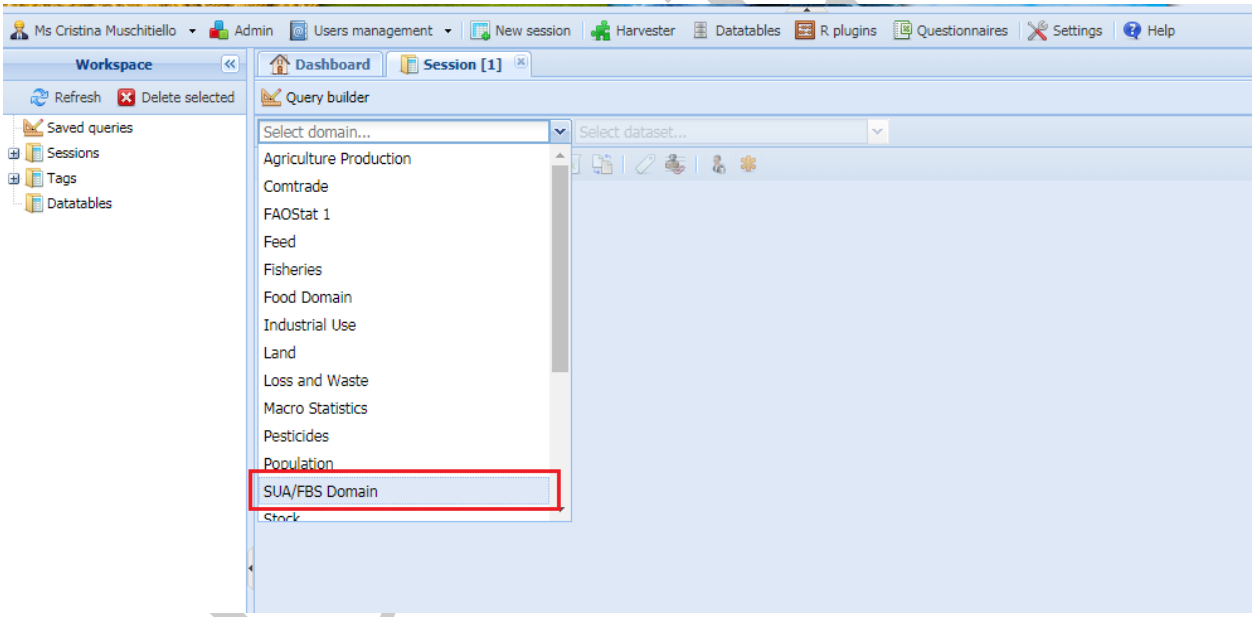


Figure 4: Select Domain

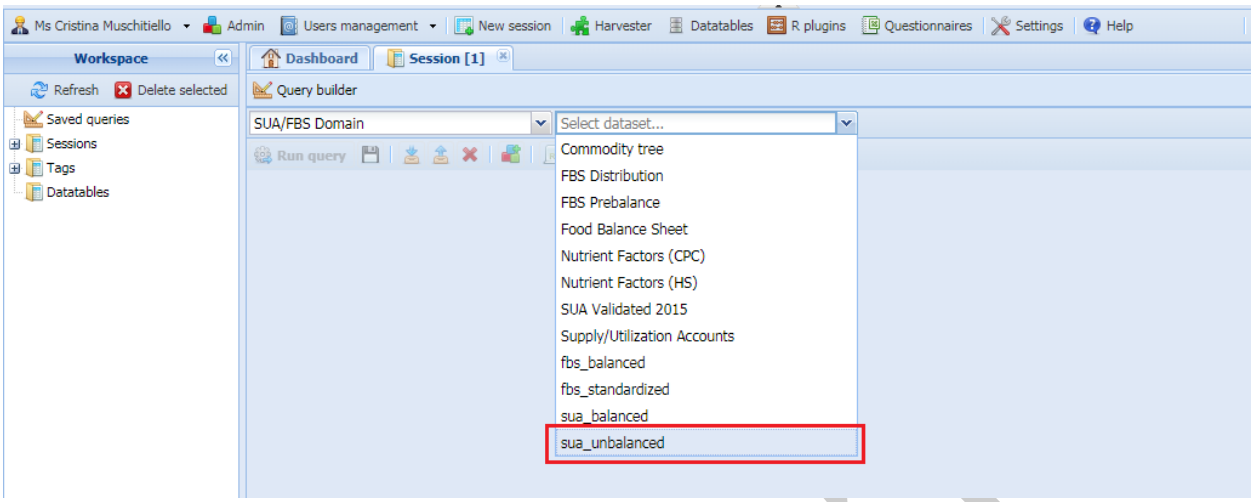


Figure 5: Select Dataset

3.1.1 Make and run the query on this session

The query has to be done only on the country for which the Pull data has to be performed. Indeed the plugin could be performed on one of the two following set of countries: *session Countries* or *all countries*. In our example *China, Mainland* is selected.

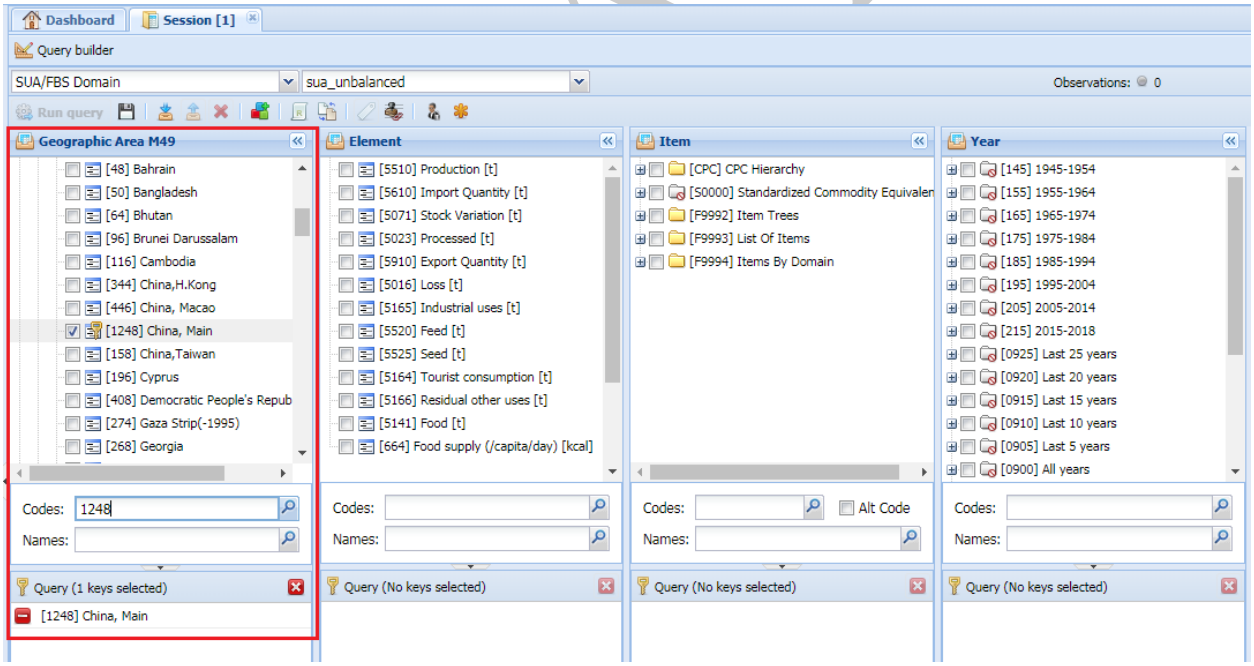


Figure 6: Select Country/ies

All elements here have to be selected (figure 7) and all items (figure 8). The years to be selected depend on the interest of the user. In this example the time range 2010:2016 is used

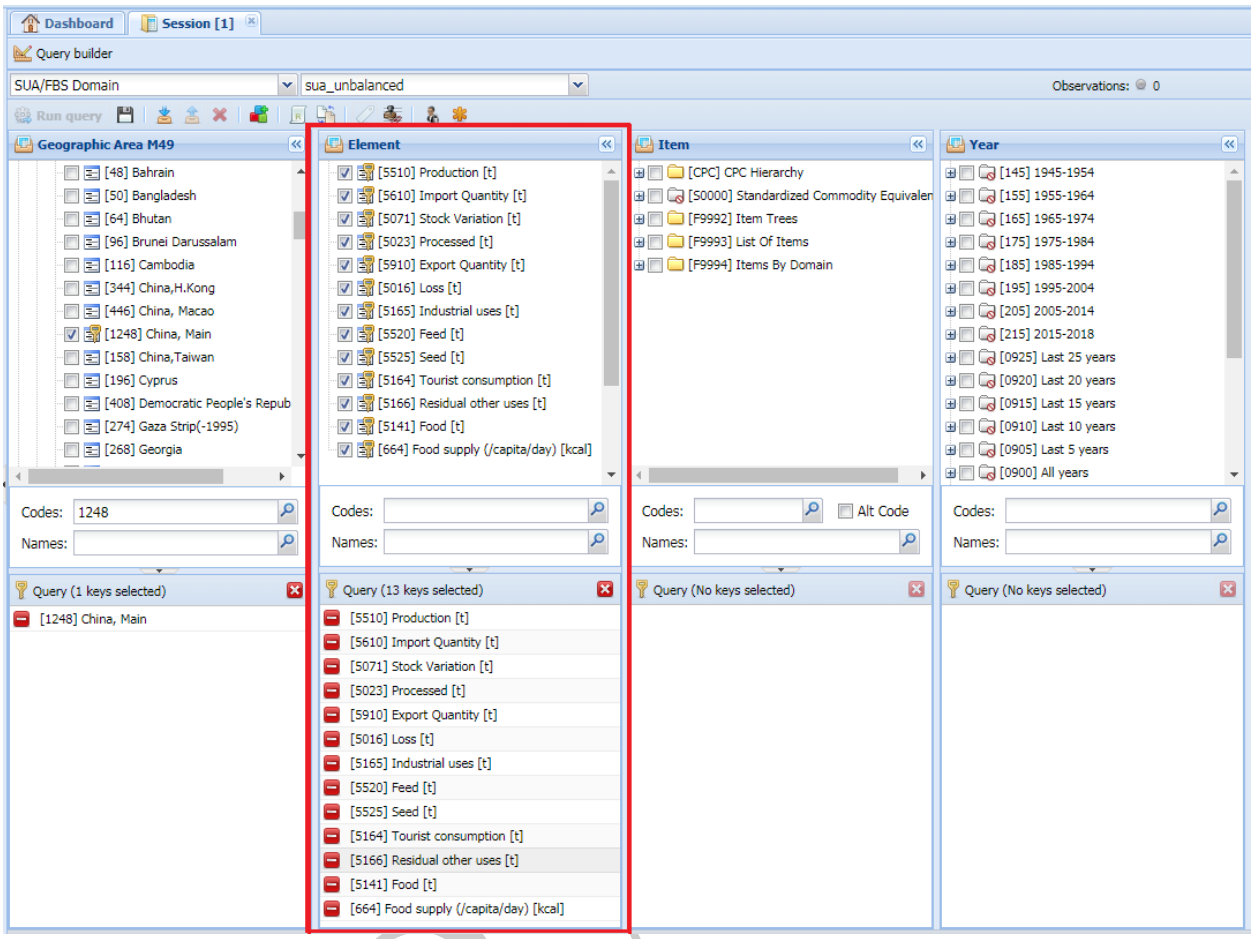


Figure 7: Select all Elements

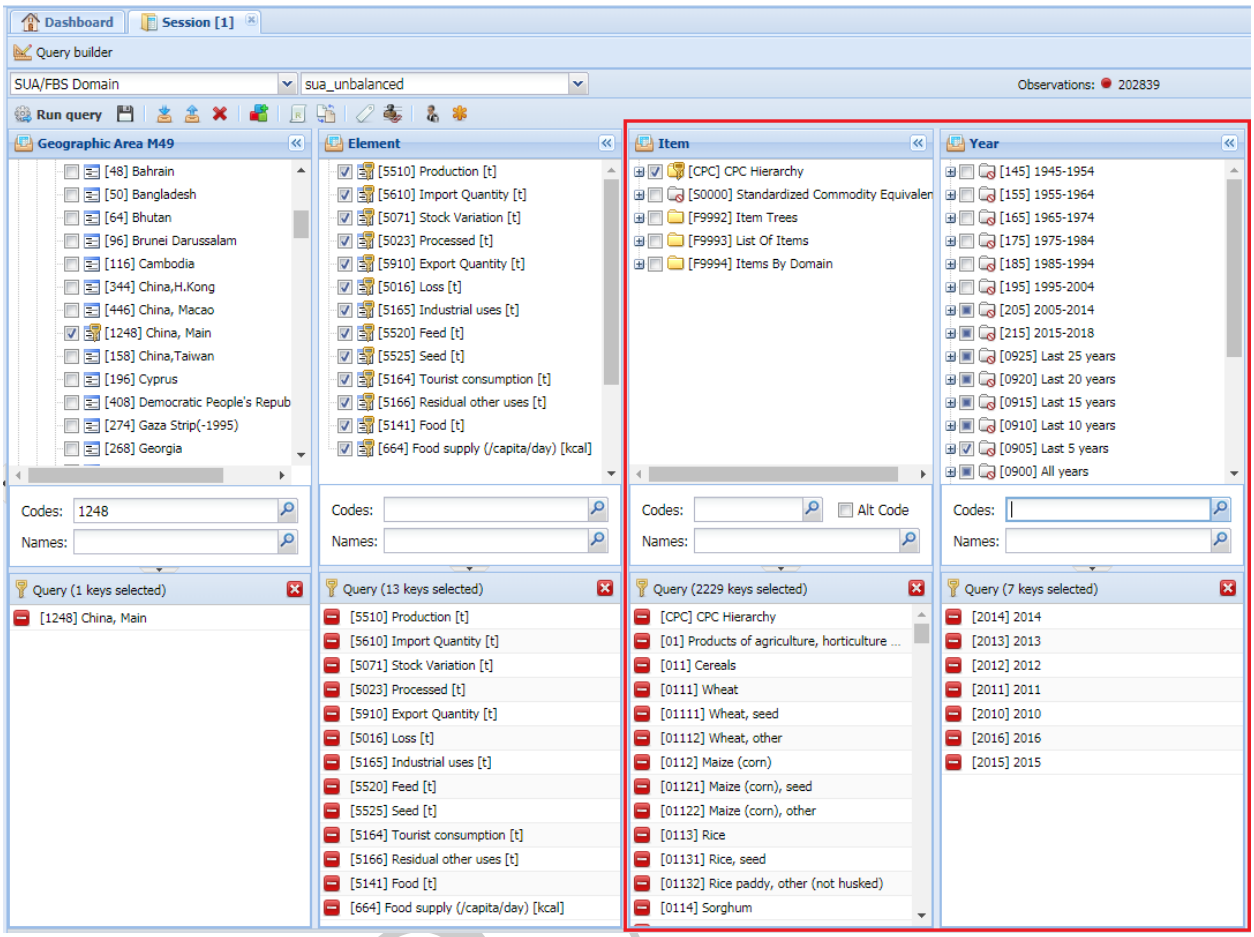


Figure 8: Select items and years

When all Variables have been defined, the query can be run:

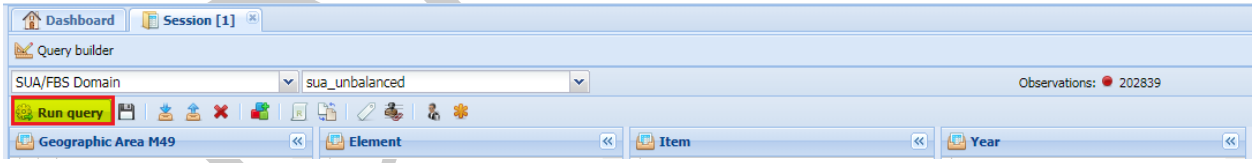


Figure 9: Run query

3.1.2 The session content

The Session just created (figure 10) contains the data that will be input of the Plugin for that country. As said in the `pullDataToSua` plug-in document[1] "*At the moment this dataset is filled with data coming from the old system (dataset "suaValidated2015"), from 2000 to 2013 for* ***all countries*** *and New data from 2014 onward*", if existing. If not existing, because the FBS have not been calculated yet, there will be blank cells to be filled.

[1] *faoswsStandardization:`pullDataToSUA` plugin*

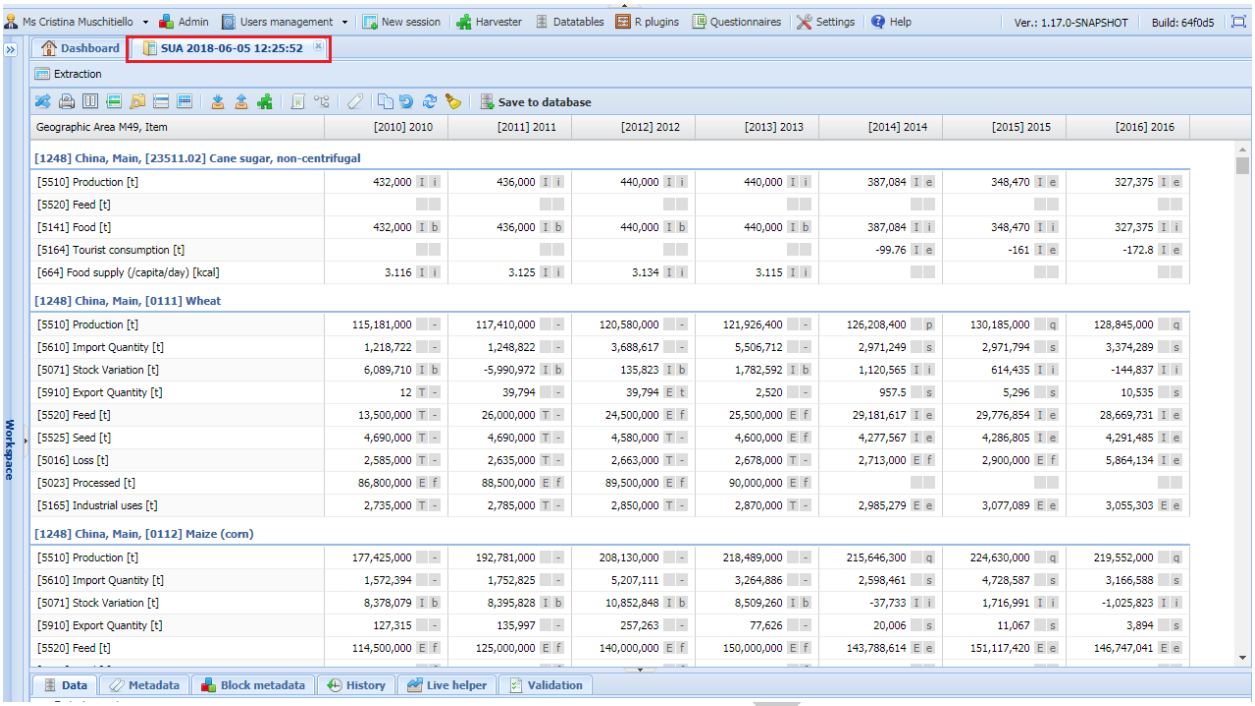


Figure 10: The Session

As previously mentioned, for the execution of the plugin and an easy managing of the operations, is better to rename the Session in a consistent and easily recognizable way.

3.1.3 Rename session

This session has the name that has been generated automatically from the SWS: *SUA 2018-06-05 12:25:52* representing the data-set, day and time of the creation of the Session. As reported in figures 11 to 13: 1. Right click on the session name 2. Select “Rename” 3. Assign a name 4. click “ok”

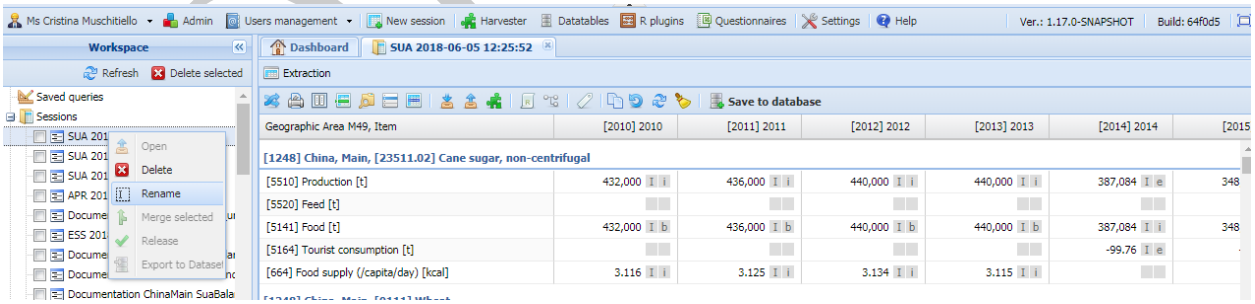


Figure 11: Rename Session - 1

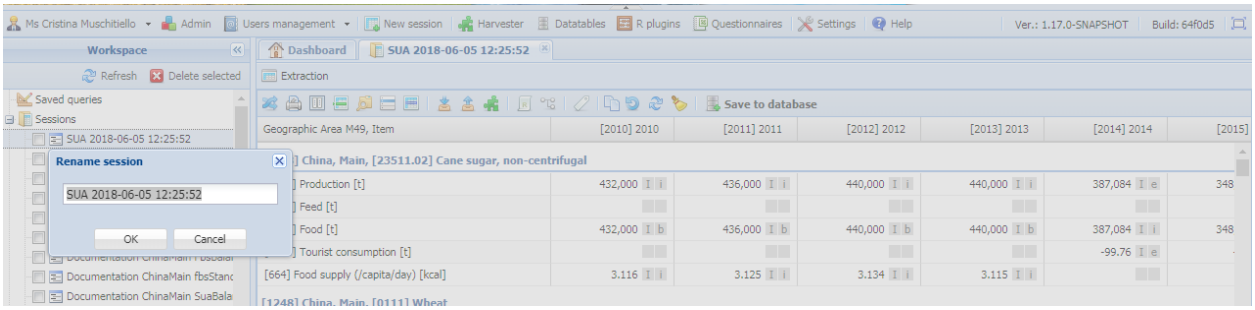


Figure 12: Rename Session - 2

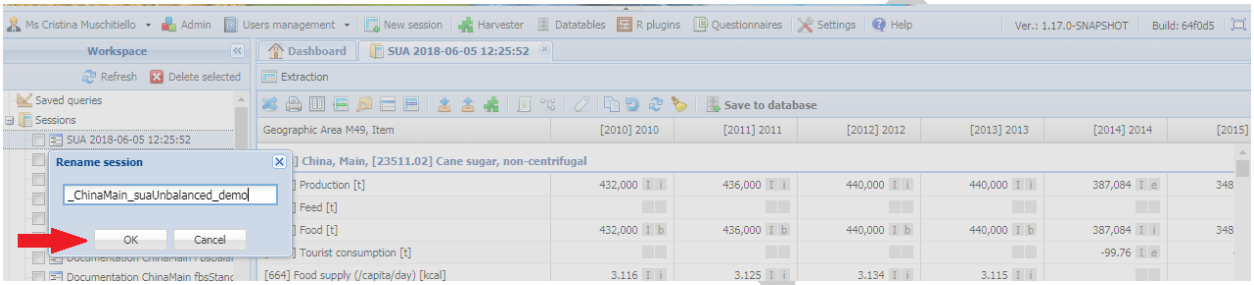


Figure 13: Rename Session - 3

3.2 `sua-balanced` session

A second session has to be opened on the domain:data-set *SUA/FBS:sua_balanced*.

3.2.1 Make and run the query on the session/Duplicate Session

This can be one in two ways. One can re-do all the steps for a new session, as reported in figures from 14 to 17 or *Duplicate* a session.

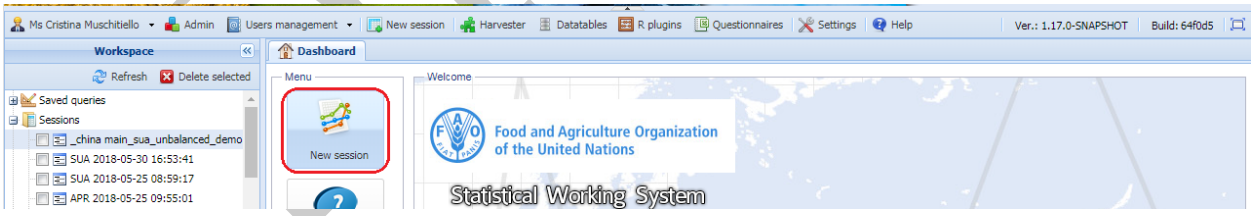


Figure 14: Sua balanced session - 1

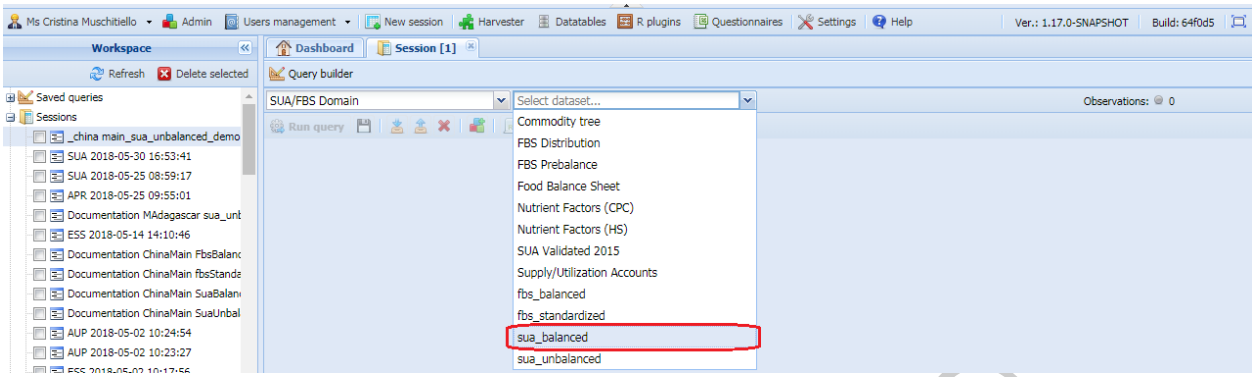


Figure 15: Sua balanced session - 2

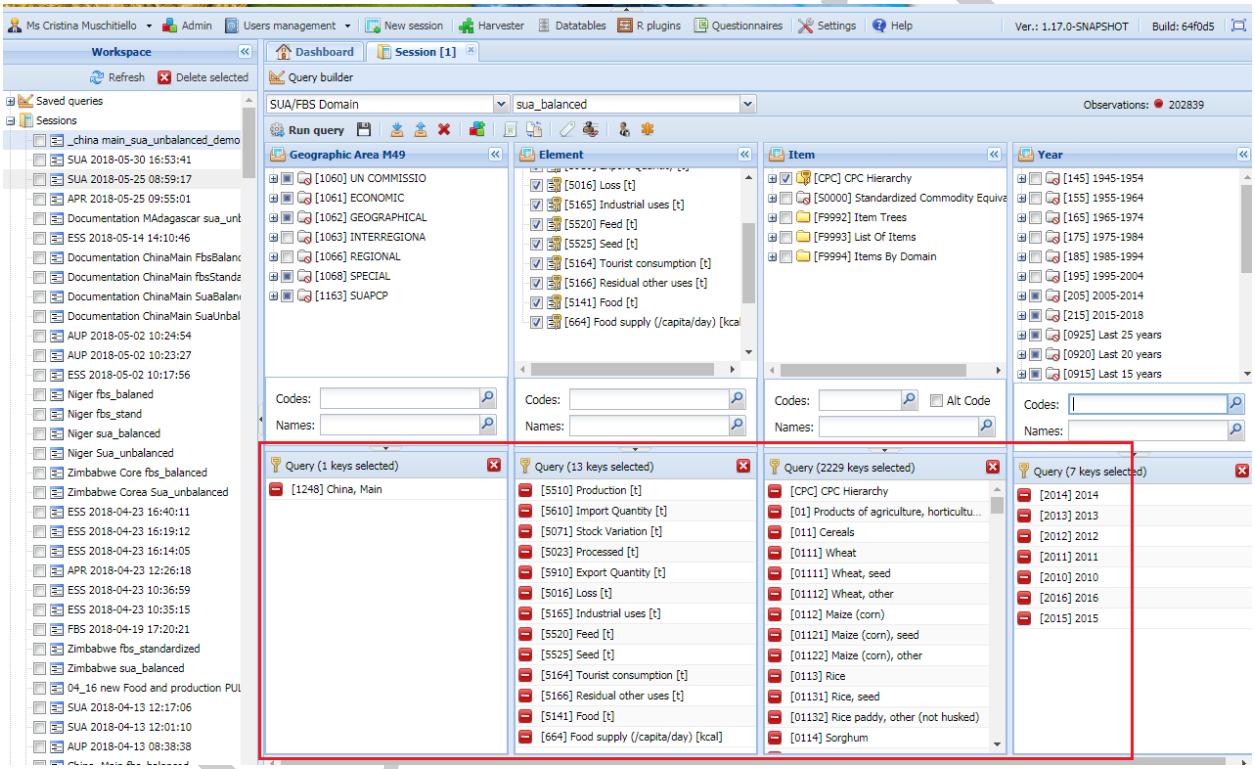


Figure 16: Sua balanced session - 3

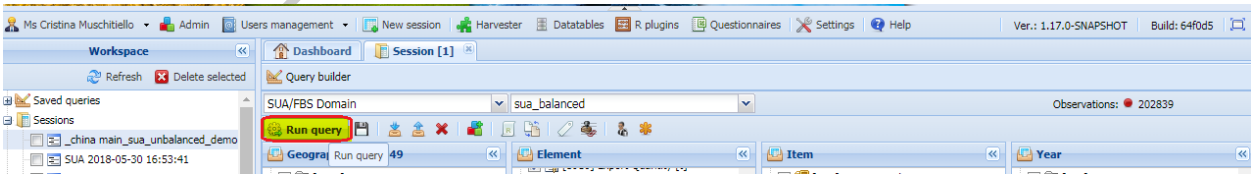


Figure 17: Sua balanced session - 4

Instead of re-doing all these step, an alternative is that of ***duplicate*** a session. Indeed, any time one want to create a session on the same data-set of another or on a different data-set but same set of data, is possible

to select the *duplicate session* button (figure 18). The *Duplicate Session* option open a new window with a pre-set query identical to the one from which the session is duplicated. This new pre-set query is still open for changes, therefore, from here is possible to change the data-set and obtain the new session without having to select again all the variables. In our example one should do the *duplicate session* on the *sua_unbalanced* table and then change the data-set to *sua_balanced* (figure 19). This would allow for saving time in creating the new session on the second data-set, just selection the desired data-set and then running the query

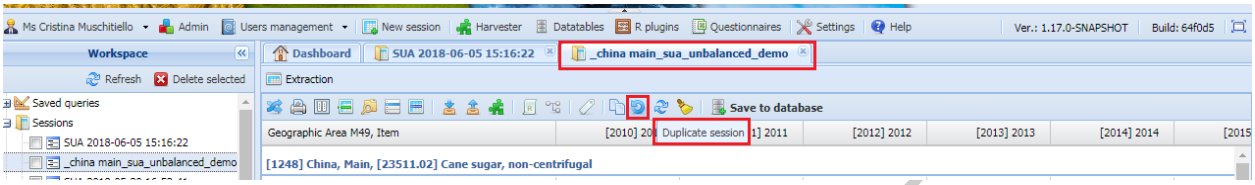


Figure 18: Duplicate Session on sua unbalanced - 1

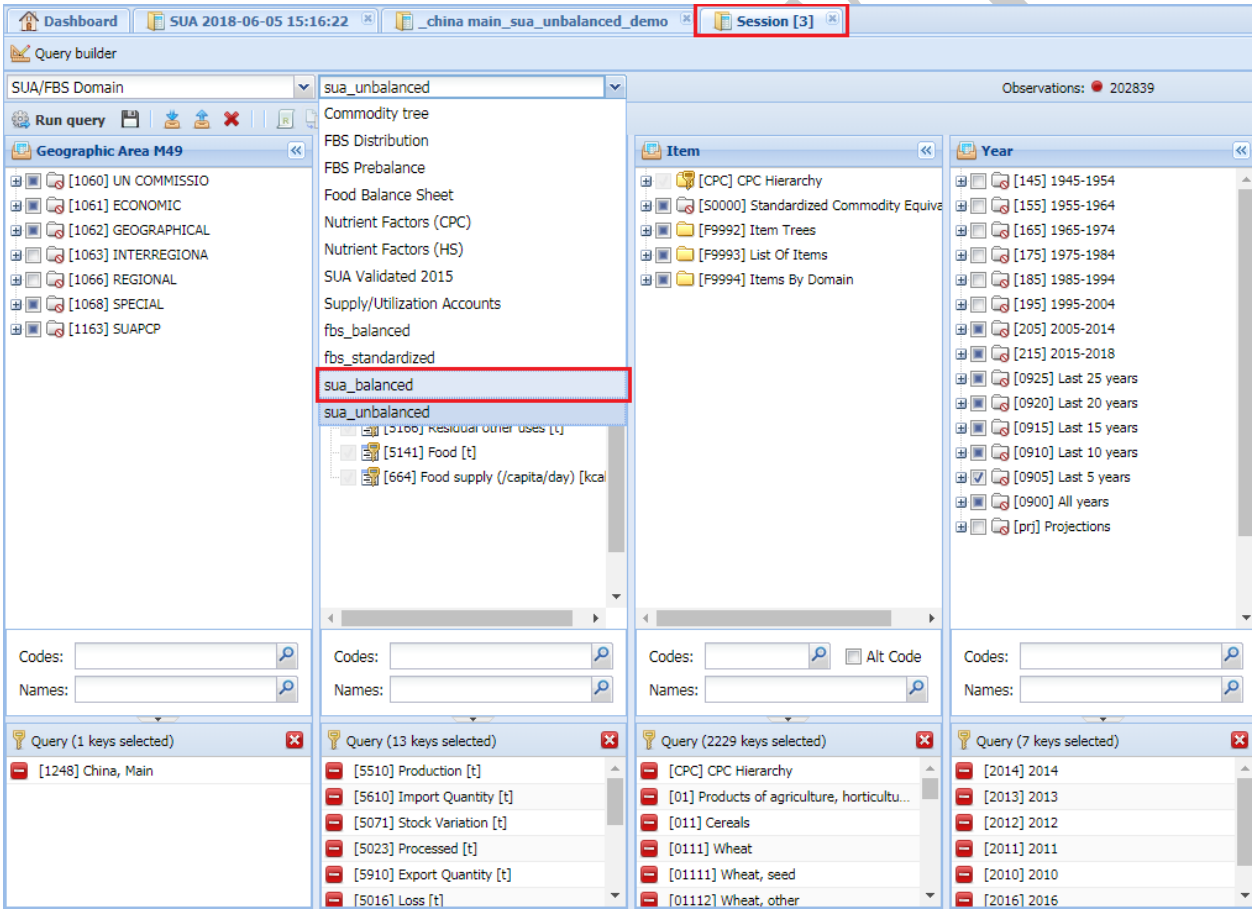


Figure 19: Duplicate Session on sua unbalanced - 2

3.2.2 Session content

Also *sua_balanced* data-set has been filled with *data coming from the old system (dataset "suaValidated2015"), from 2000 to 2013 for* ***all countries*** *and New data from 2014 onward.*

3.2.3 Rename session

Also this session has to be renamed (figure 20).

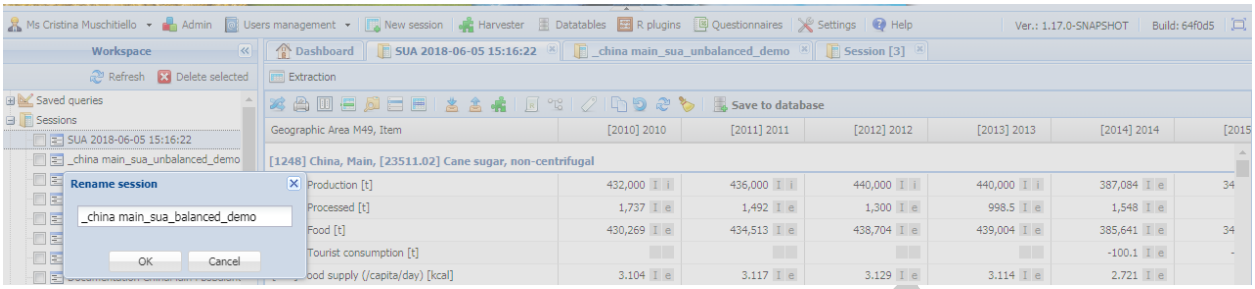


Figure 20: Rename session sua balanced

3.3 `fbs-standardized` session

This session is created using exactly the same steps just explained for the previous one. After the execution, the session has to be renamed.

3.3.1 Make and run the query on the session/Duplicate Session

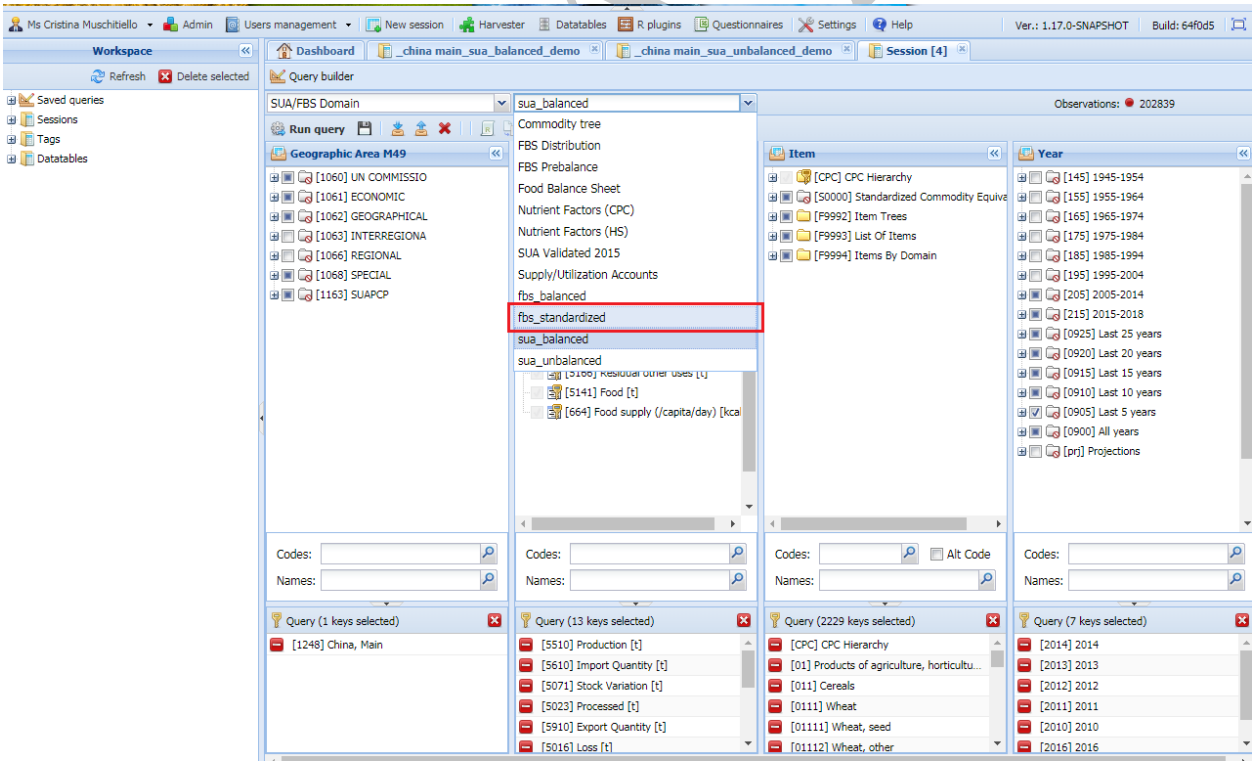


Figure 21: Duplicate balanced session in the fbs standardized dataset

3.3.2 Session content

Because the old system did no have this intermediate step and there was no old data stored to copy, this data-set has blank valued up to 2013 and new data from 2014. The new data are available for countries that have been already processed for FBS.

3.3.3 Rename session

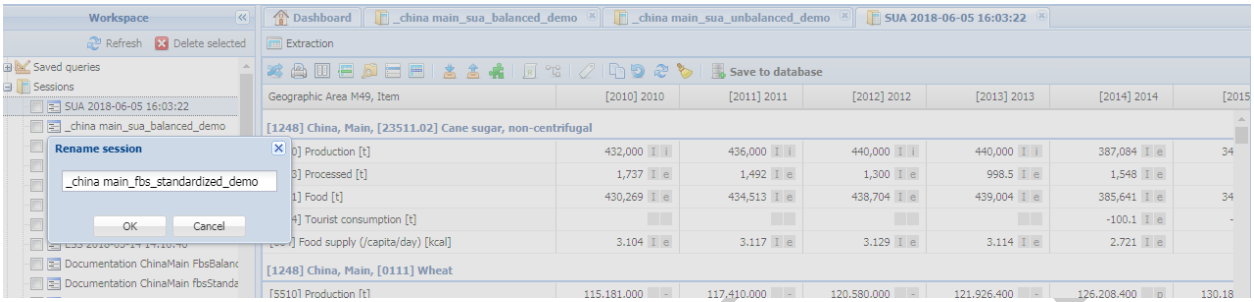


Figure 22: Rename session fbs standardized

3.4 `fbs-balanced` session

This is the main output data-set.

3.4.1 Make and run the query on the session/Duplicate Session

In the use of the *duplicate session* option all the FBS item have to be selected in addition to the CPC, because this data-set do not contain CPC item, therefore the session would come empty[2]. Also nutritive factors have to be selected in this step (figure 23)

[2]this is just a visualization need, in the sense that even if the session is empty, it would be filled anyway after the plug-in will be run. Anyhow, the previous years would not be visualized if the FBS item are not selected

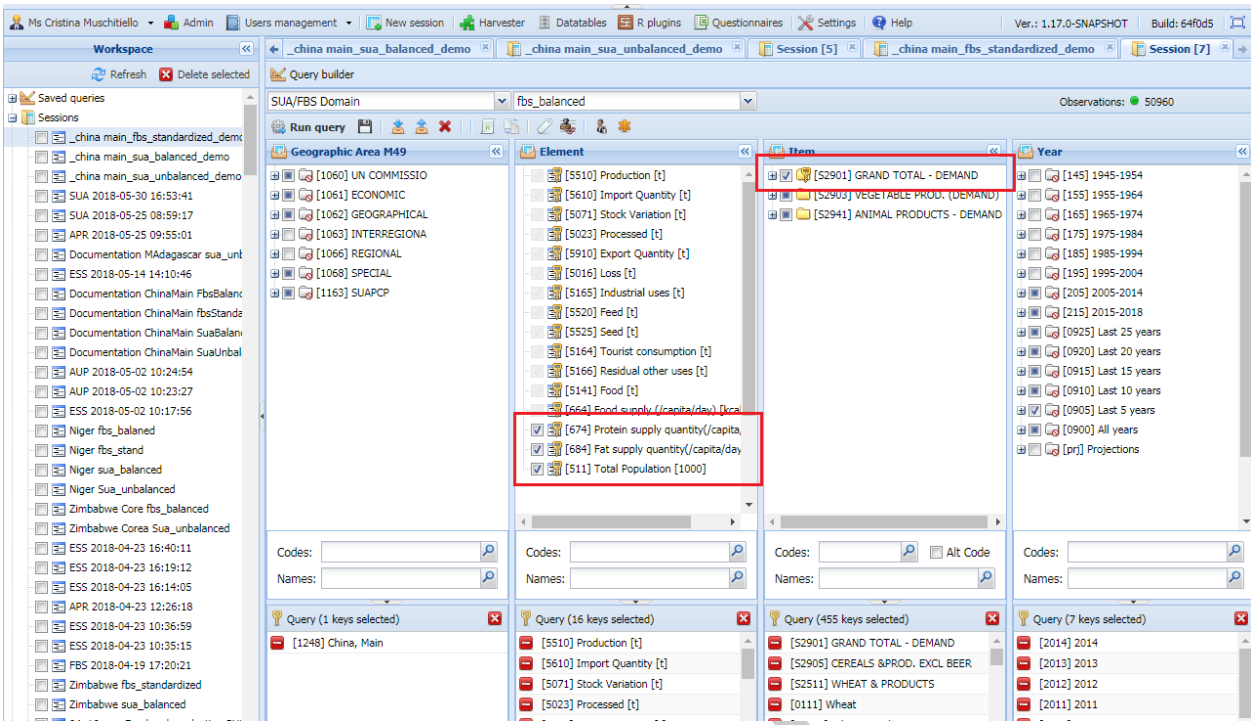


Figure 23: Dupicate fbs Standardized

3.4.2 Session content

FBS data coming from the old System are stored here from 2000 to 2013, while new FBS data are stored from 2014 onward. The new data are available for countries that have been already processed for FBS.

3.4.3 Rename session

Rename as the other data-sets.

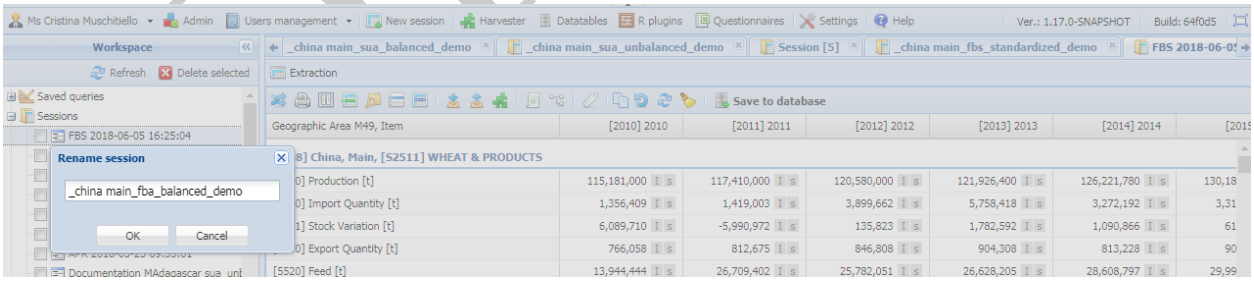


Figure 24: Rename fbs balanced

4 Select plug-in

In the plug-in window, elect the `Full Standardization and Balancing` Plug-in. This opens the window in figure 27.

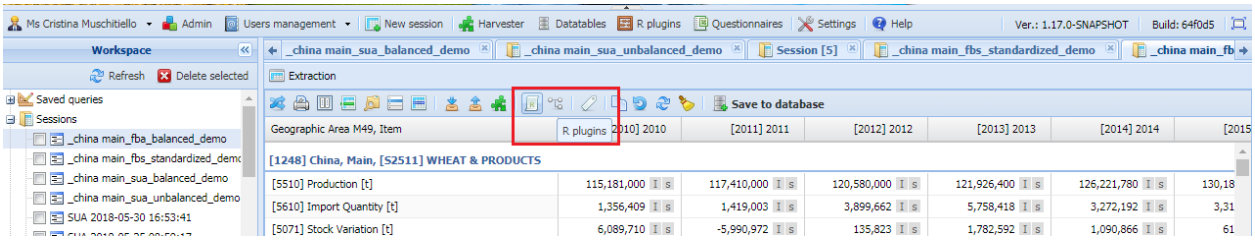


Figure 25: Select plug-in

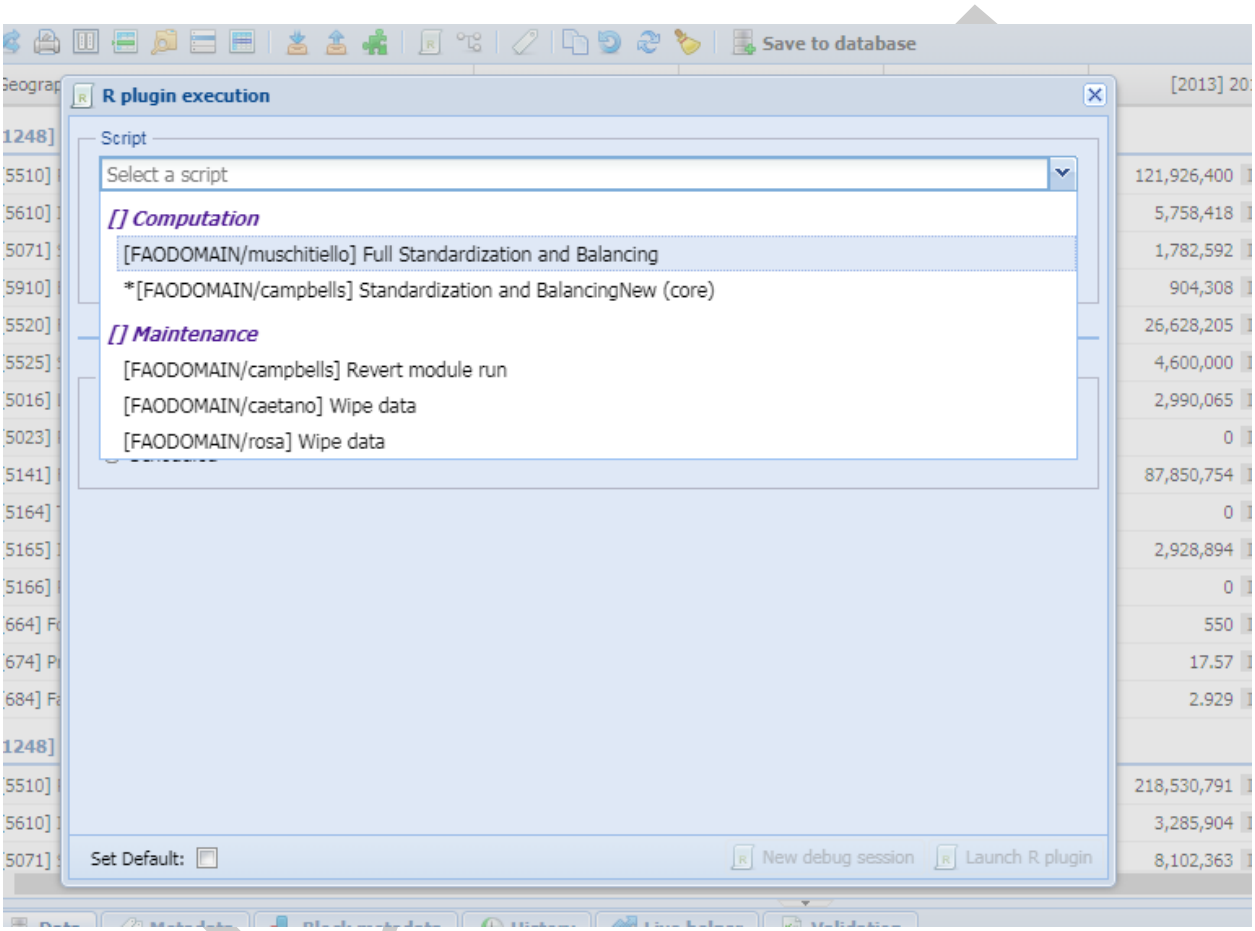


Figure 26: Plug-in window

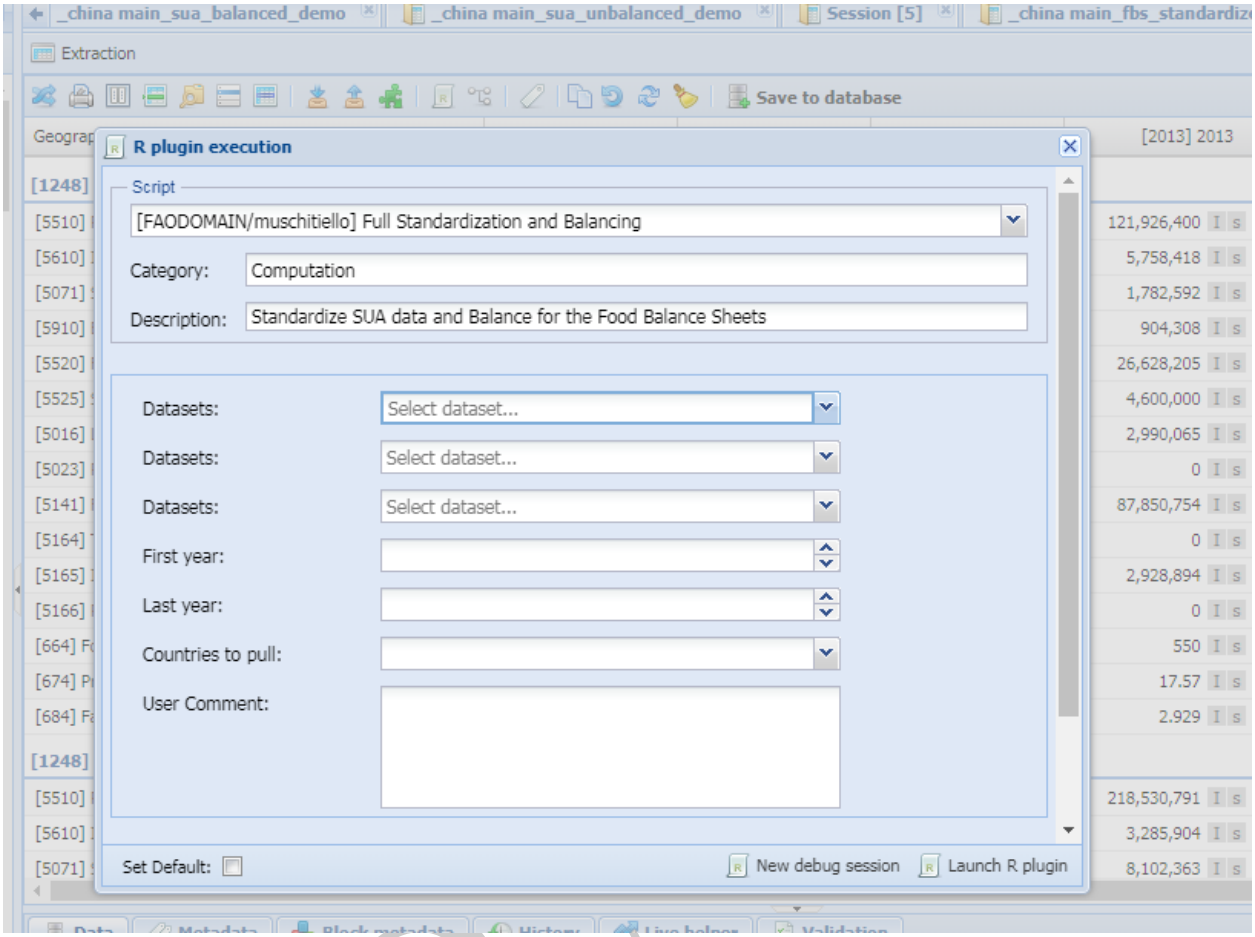


Figure 27: Plug-in parameters - 1

There are 3 *Dataset* sections. These are made for specifying the sessions in which output data have to be saved. The name of the data-set is reported in the fist line, while the name of the sessions are in the following lines (figure 28). From the drop-down menu, elect the session you are working in (figures 28 and 29).

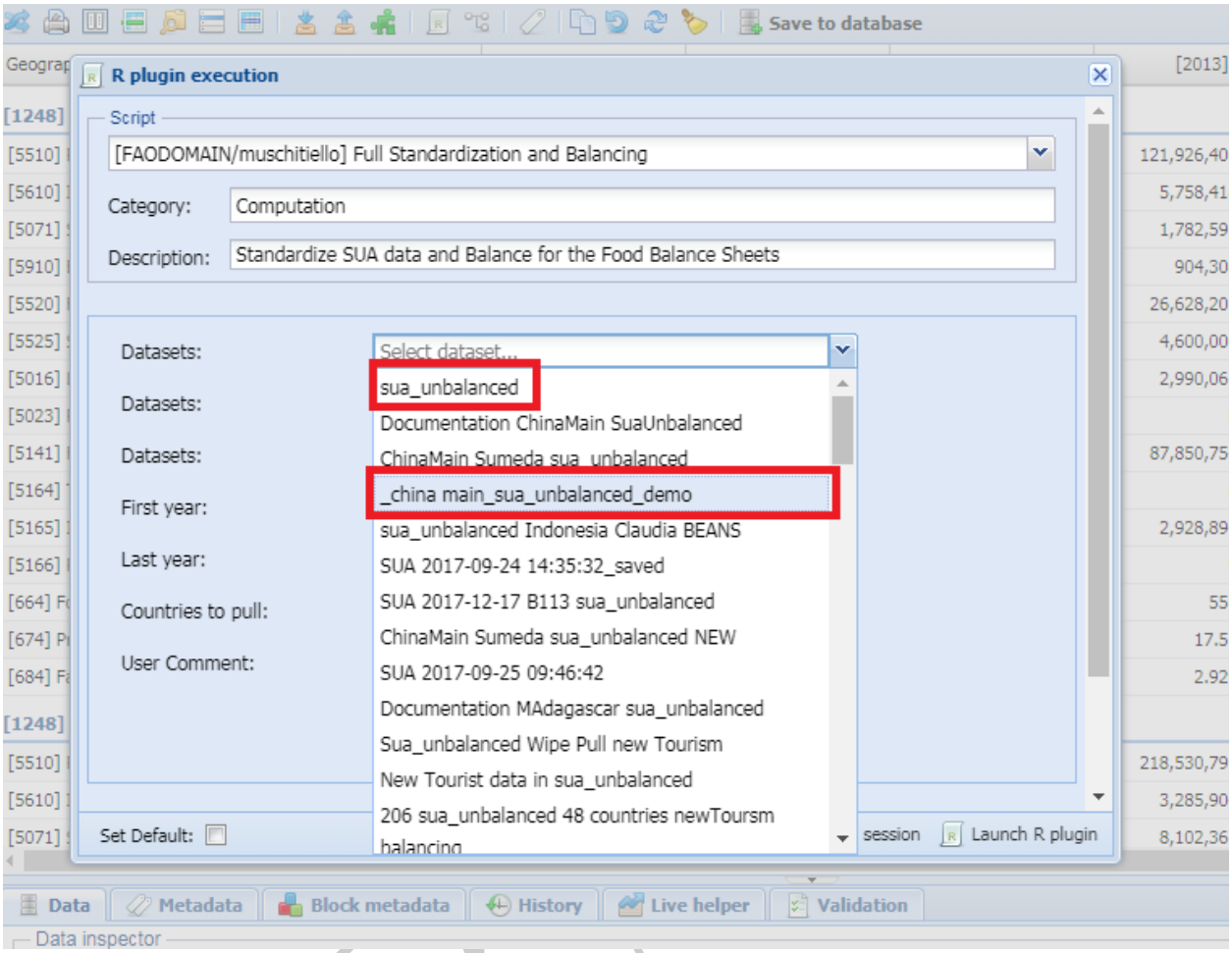


Figure 28: Plug-in parameters - 2

Also the time range has to be selected here, which may or may not be the same as the time range of the session. In this example, the session has the time range 2010-2016 but the standardization is run for the period 2014-2016.

Finally, also the countries have to be selected. In particular, is possible to select *all countries* or *session countries* (figure 29).

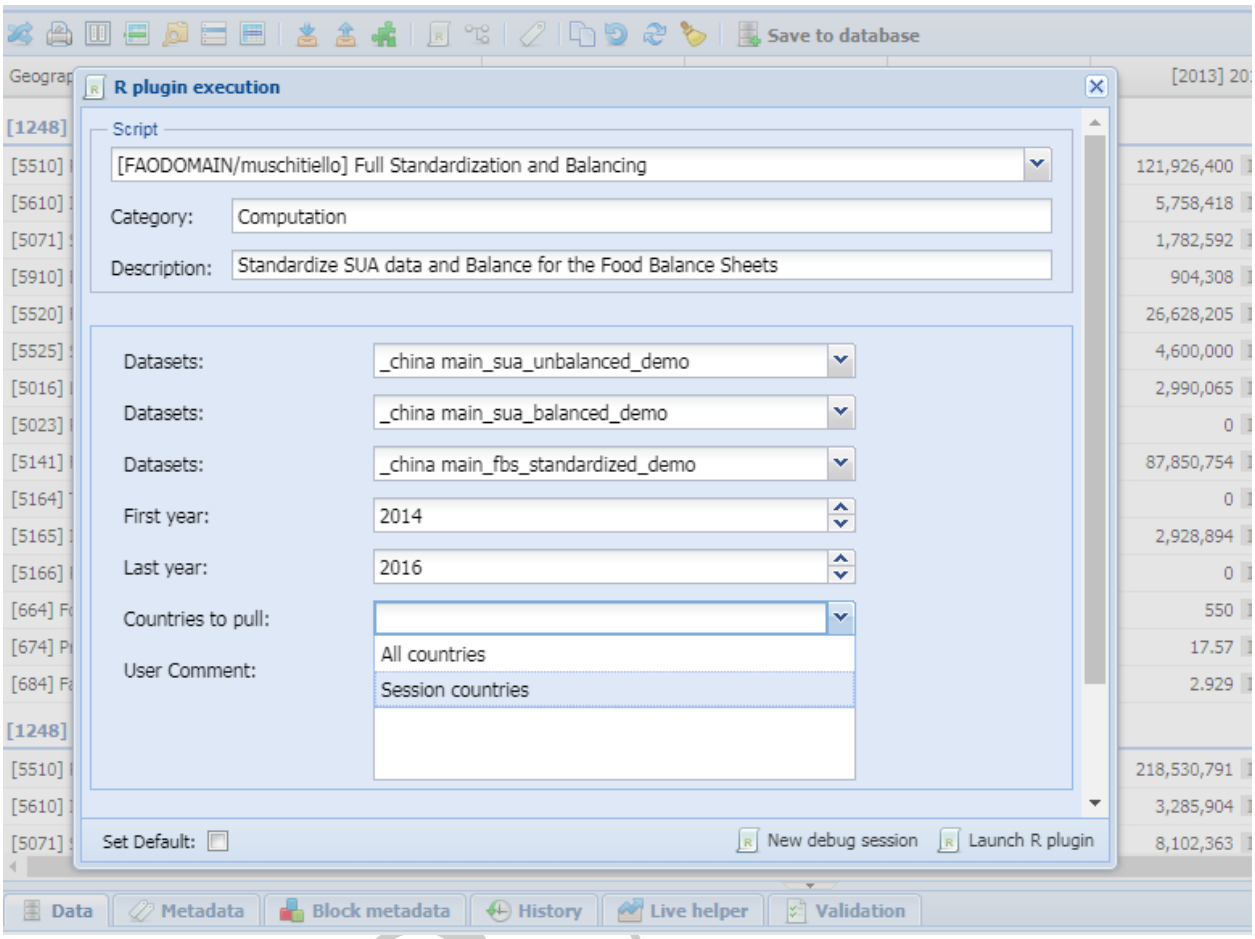


Figure 29: Plug-in parameters - 3

After all parameters are set, just launch the plug-in (figure 30).

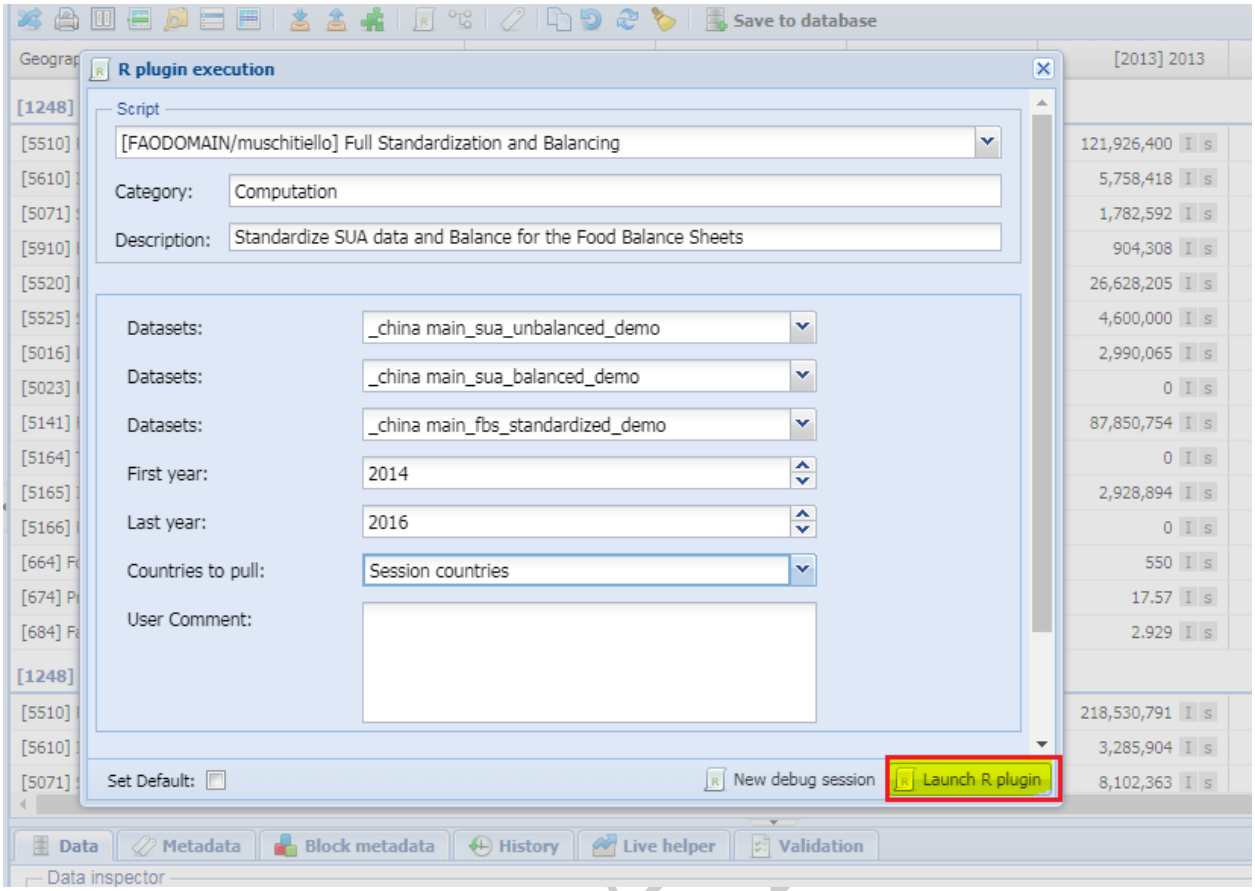


Figure 30: Launch Plug-in

5 Plug-in Run

While the plug-in runs, some email are sent to the user who launches it:

1. Commodity tree download and check

   the first email tells that the Commodity tree has been correclty downoaded. Moreover it also tells that the extraction rates are valid. This message is sent because the standardization plug-in performs a validity check on extraction rates. If extraction rates are outside some specified range or if there are wrong flag combinations, the process stops. This validation is needed because both eR and Shares might be manually changed from any user. If some error is introduced, it has to be checked otherwise FBS might have inconsistent or wrong results. The first validation only check extraction Rates. For this reason the message says taht also a check on shares is needed.

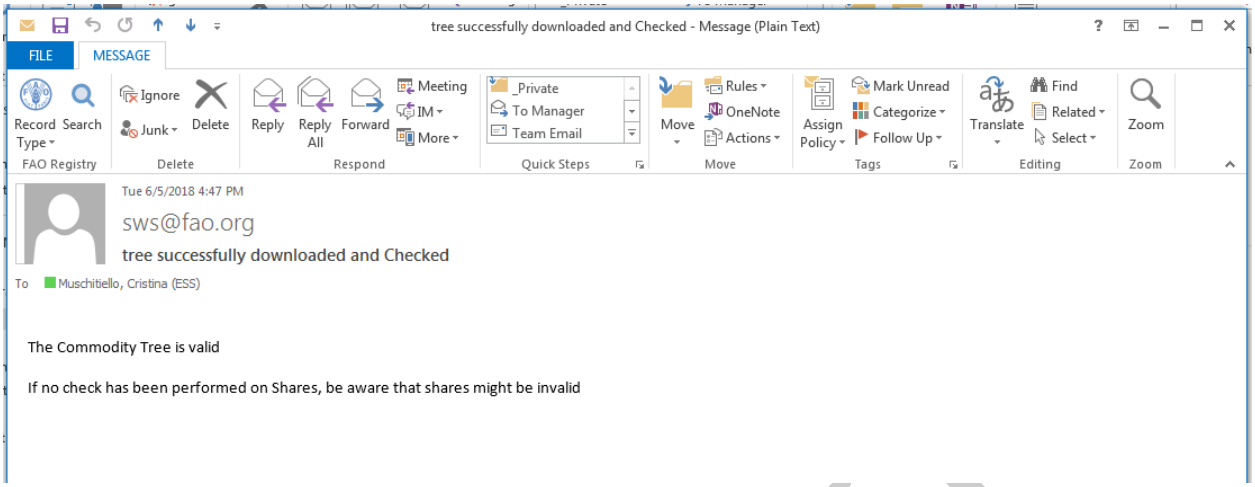


Figure 31: Tree Validation email - 1

2. Check on shares

   Shares also migh be manually changed by any user. A second check is introduce for shares. The validation function checks if shares sum up to 1 by child. If the check runs correctly, the email is sent (figure 32), if not, the plug-in stops running. The two checks are kept separated because the R-functions that perform these validity cheks might be separatedly needed by other modules and plug-ins[3].

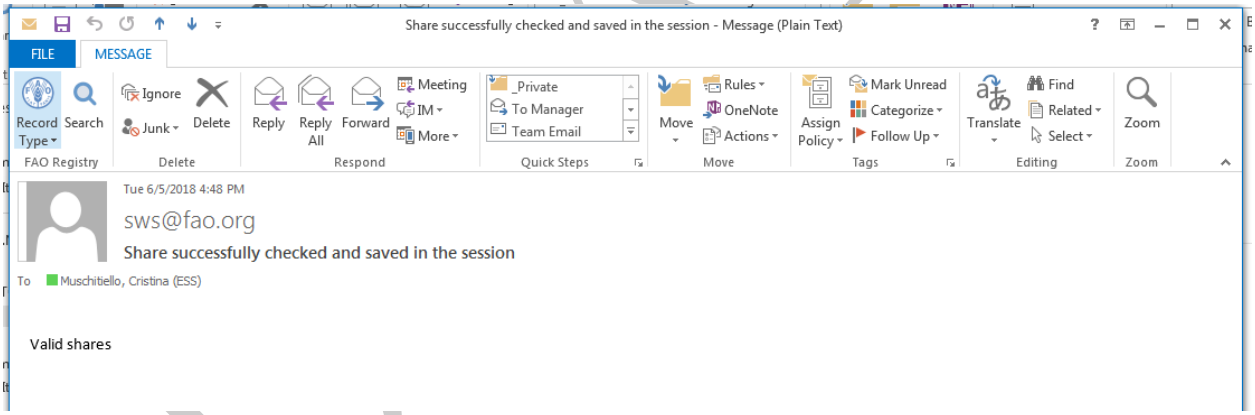


Figure 32: Tree Validation email - 2

3. The plugin runs correctly BUT SOME PRODUCTION FIGURES HAVE BEEN CHANGED during the SUA Filling.

   As discussed, it might happen that, if Utilization > supplies and it is not enough to reduce utilizations for balancing one SUA line, the module DO CHANGE OFFICIAL PRODUCTION FIGURES. In this case, the changed figures are saved in a csv file and sent by email. This email represents a Warning: Those SUA have to be double-checked and corrected. An example of the csv file follows. That file contains the entire SUA in which the production value is changed, by Country/commodity/year combination. This file is only for visualization. The changes have to be performed directly in the SWS.

4. Final Email

[3]Developers always tries to keep functions general

The plug-in may require a lomg time to run, depending on how many countries are in the session and, also, from the sWS congestion. As a consequece, the session may expire in the meantime, making it difficult to understand when the plug-in ends running. The sending of an email tells when the plug-in has run correctly (34). If the user is still on the session when it happens, a message is, also visualized on the screen (33).

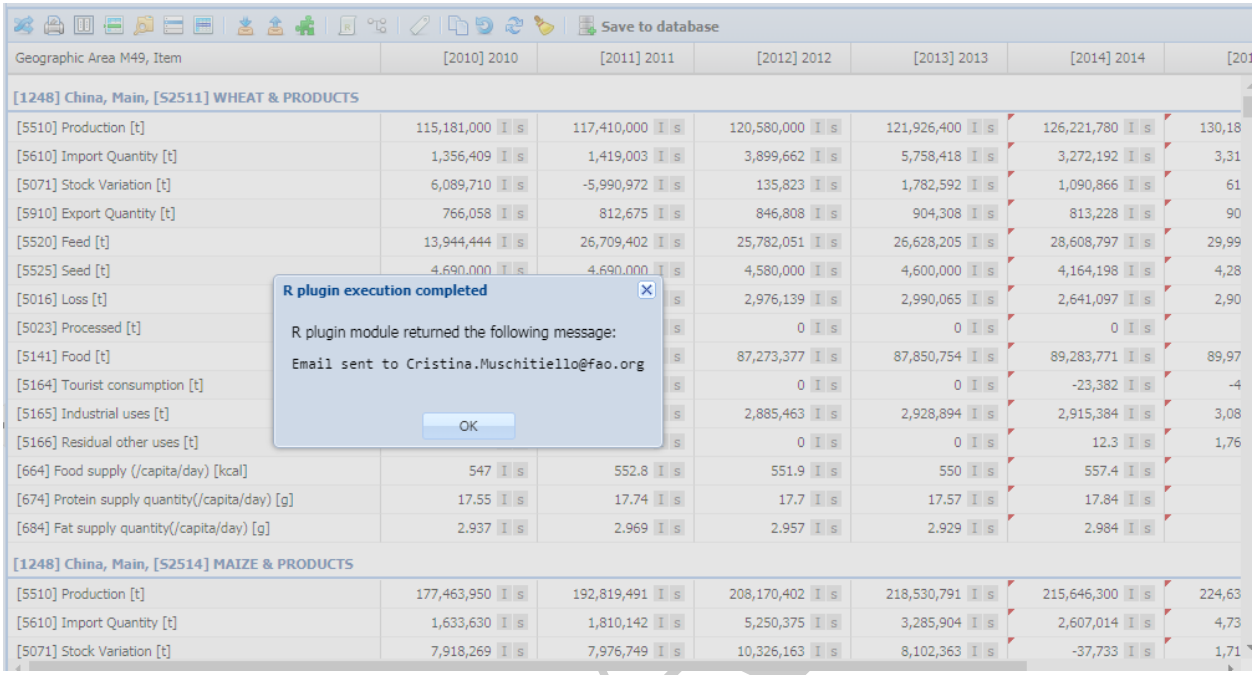


Figure 33: Run message

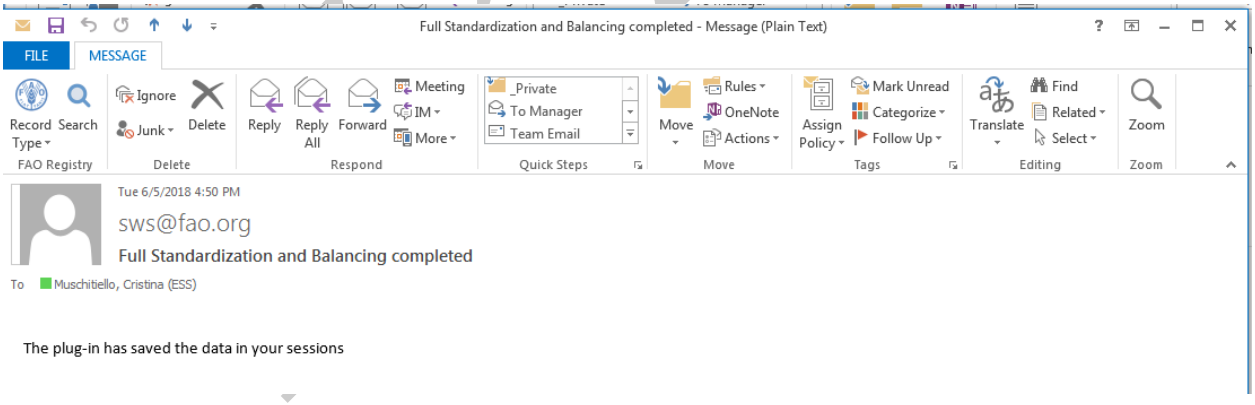


Figure 34: Final email

6 The sessions after saving

6.1 fbs_balanced session

Values have changed IN THE TIME RANGE FOR WHICH THE PLUGIN HAS RUN (2014:2016), therefore red signs will appear in the cell of the selected time range. In our example, nothing changes in the time range

2010-2013 (35).

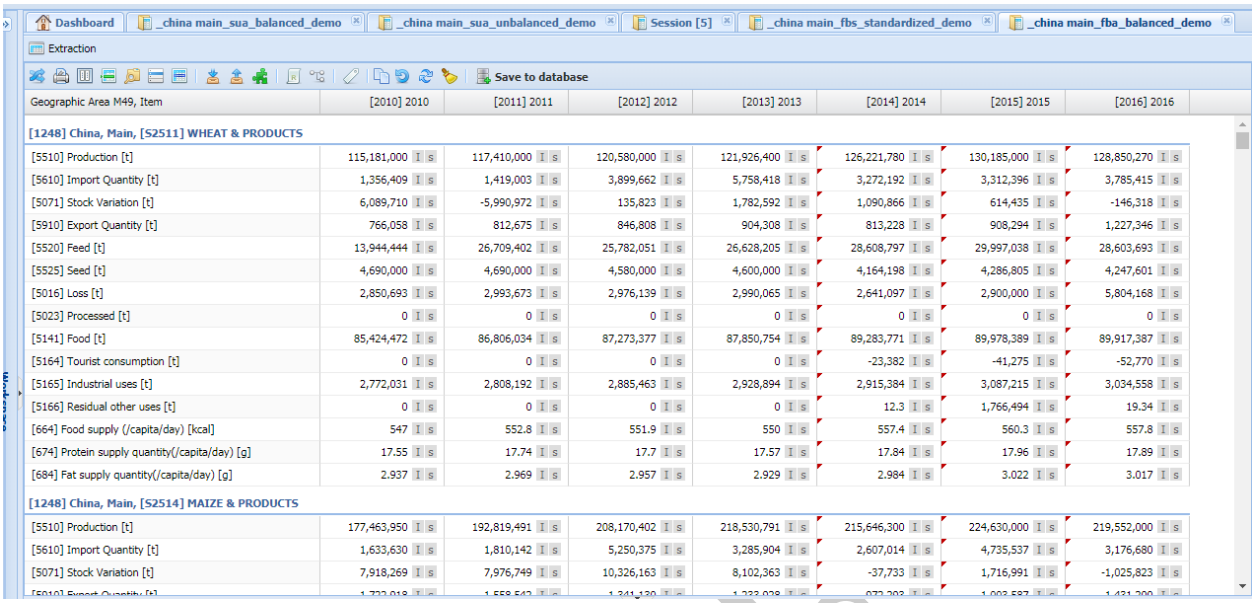


Figure 35: The session after the Run

6.2 fbs_standardized session

Values have changed IN THE TIME RANGE FOR WHICH THE PLUGIN HAS RUN (2014:2016), therefore red signs will appear in the cell of the selected time range. In our example, nothing changes in the time range 2010-2013 (36).

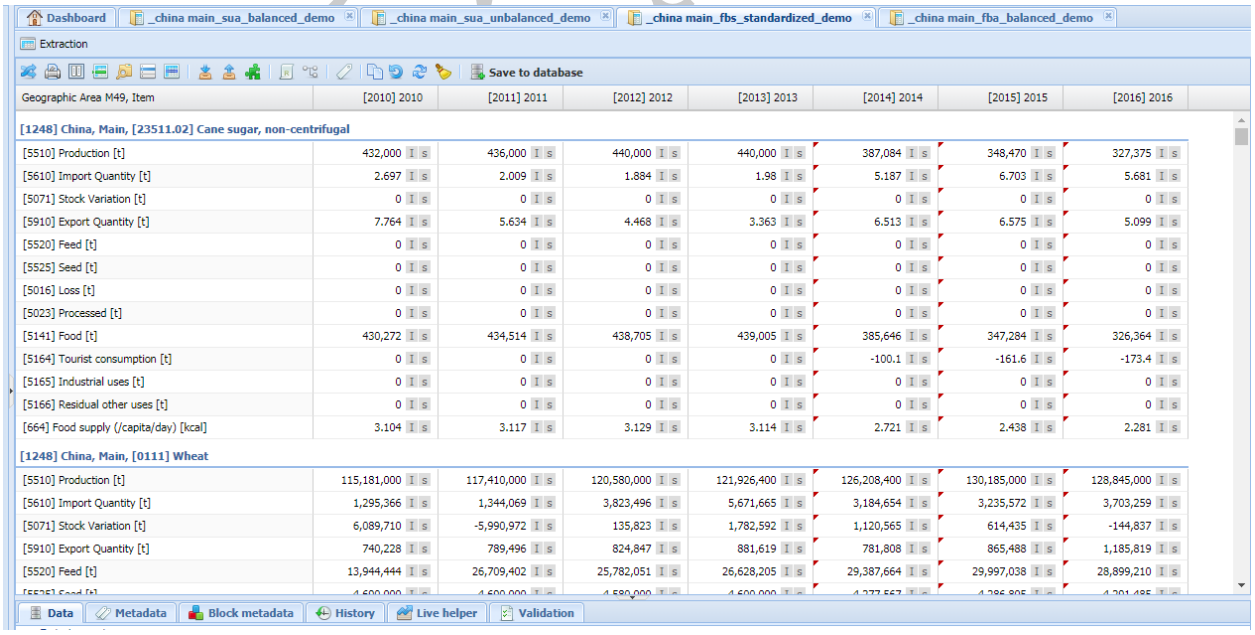
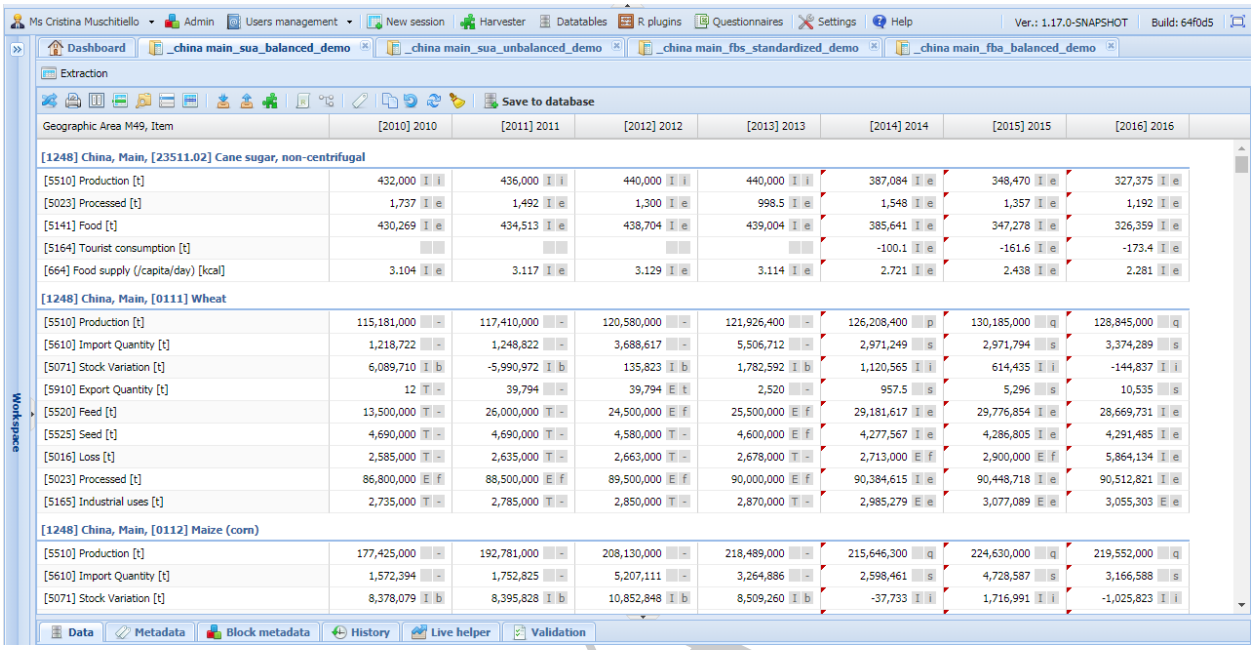


Figure 36: fbs standardized after the plug-in run

6.3 sua_balanced session

Here values have changed IN THE TIME RANGE FOR WHICH THE PLUGIN HAS RUN (2014:2016), therefore red signs will appear in the cell of the selected time range. Nothing is changed in the other time ranges, instead: the same data as sua_unbalanced are shown for years up to 2013.



Geographic Area M49, Item	[2010] 2010	[2011] 2011	[2012] 2012	[2013] 2013	[2014] 2014	[2015] 2015	[2016] 2016
[1248] China, Main, [23511.02] Cane sugar, non-centrifugal							
[5510] Production [t]	432,000 I i	436,000 I i	440,000 I i	440,000 I i	387,084 I e	348,470 I e	327,375 I e
[5023] Processed [t]	1,737 I e	1,492 I e	1,300 I e	998.5 I e	1,548 I e	1,357 I e	1,192 I e
[5141] Food [t]	430,269 I e	434,513 I e	438,704 I e	439,004 I e	385,641 I e	347,278 I e	326,359 I e
[5164] Tourist consumption [t]					-100.1 I e	-161.6 I e	-173.4 I e
[664] Food supply (/capita/day) [kcal]	3.104 I e	3.117 I e	3.129 I e	3.114 I e	2.721 I e	2.438 I e	2.281 I e
[1248] China, Main, [0111] Wheat							
[5510] Production [t]	115,181,000 -	117,410,000 -	120,580,000 -	121,926,400 -	126,208,400 p	130,185,000 q	128,845,000 q
[5610] Import Quantity [t]	1,218,722 -	1,248,822 -	3,688,617 -	5,506,712 -	2,971,249 s	2,971,794 s	3,374,289 s
[5071] Stock Variation [t]	6,089,710 I b	-5,990,972 I b	135,823 I b	1,782,592 I b	1,120,565 I i	614,435 I i	-144,837 I i
[5910] Export Quantity [t]	12 T -	39,794 -	39,794 E t	2,520 -	957.5 s	5,296 s	10,535 s
[5520] Feed [t]	13,500,000 T -	26,000,000 T -	24,500,000 E f	25,500,000 E f	29,181,617 I e	29,776,854 I e	28,669,731 I e
[5525] Seed [t]	4,690,000 T -	4,690,000 T -	4,580,000 T -	4,600,000 E f	4,277,567 I e	4,286,805 I e	4,291,485 I e
[5016] Loss [t]	2,585,000 T -	2,635,000 T -	2,663,000 T -	2,678,000 T -	2,713,000 E f	2,900,000 E f	5,864,134 I e
[5023] Processed [t]	86,800,000 E f	88,500,000 E f	89,500,000 E f	90,000,000 E f	90,384,615 I e	90,448,718 I e	90,512,821 I e
[5165] Industrial uses [t]	2,735,000 T -	2,785,000 T -	2,850,000 T -	2,870,000 T -	2,985,279 E e	3,077,089 E e	3,055,303 E e
[1248] China, Main, [0112] Maize (corn)							
[5510] Production [t]	177,425,000 -	192,781,000 -	208,130,000 -	218,489,000 -	215,646,300 q	224,630,000 q	219,552,000 q
[5610] Import Quantity [t]	1,572,394 -	1,752,825 -	5,207,111 -	3,264,886 -	2,598,461 s	4,728,587 s	3,166,588 s
[5071] Stock Variation [t]	8,378,079 I b	8,395,828 I b	10,852,848 I b	8,509,260 I b	-37,733 I i	1,716,991 I i	-1,025,823 I i

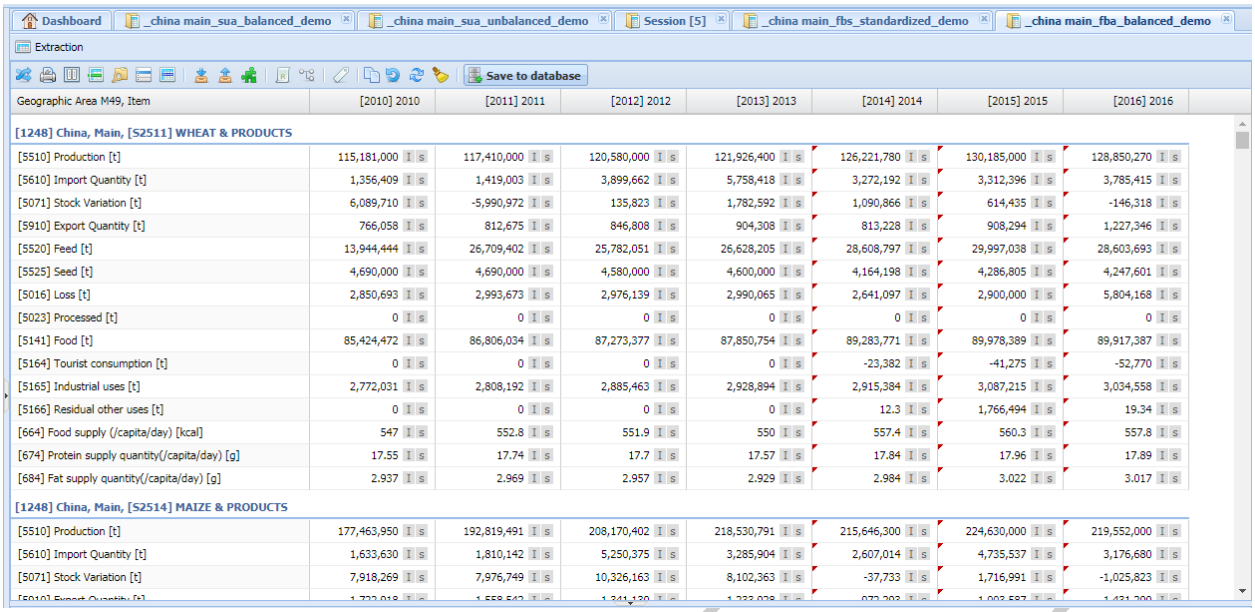
Figure 37: sua balanced after the plug-in run

6.4 sua_unbalanced session

This is the input data-set, therefore no changes in values happen in this table.

7 Final Save into the SWS

For the new figures to be used in following steps, the data have to be saved to the database.



Geographic Area M49, Item	[2010] 2010	[2011] 2011	[2012] 2012	[2013] 2013	[2014] 2014	[2015] 2015	[2016] 2016
[1248] China, Main, [S2511] WHEAT & PRODUCTS							
[5510] Production [t]	115,181,000	117,410,000	120,580,000	121,926,400	126,221,780	130,185,000	128,850,270
[5610] Import Quantity [t]	1,356,409	1,419,003	3,899,662	5,758,418	3,272,192	3,312,396	3,785,415
[5071] Stock Variation [t]	6,089,710	-5,990,972	135,823	1,782,592	1,090,866	614,435	-146,318
[5910] Export Quantity [t]	766,058	812,675	846,808	904,308	813,228	908,294	1,227,346
[5520] Feed [t]	13,944,444	26,709,402	25,782,051	26,628,205	28,608,797	29,997,038	28,603,693
[5525] Seed [t]	4,690,000	4,690,000	4,580,000	4,600,000	4,164,198	4,286,805	4,247,601
[5016] Loss [t]	2,850,693	2,993,673	2,976,139	2,990,065	2,641,097	2,900,000	5,804,168
[5023] Processed [t]	0	0	0	0	0	0	0
[5141] Food [t]	85,424,472	86,806,034	87,273,377	87,850,754	89,283,771	89,978,389	89,917,387
[5164] Tourist consumption [t]	0	0	0	0	-23,382	-41,275	-52,770
[5165] Industrial uses [t]	2,772,031	2,808,192	2,885,463	2,928,894	2,915,384	3,087,215	3,034,558
[5166] Residual other uses [t]	0	0	0	0	12.3	1,766,494	19.34
[664] Food supply (/capita/day) [kcal]	547	552.8	551.9	550	557.4	560.3	557.8
[674] Protein supply quantity(/capita/day) [g]	17.55	17.74	17.7	17.57	17.84	17.96	17.89
[684] Fat supply quantity(/capita/day) [g]	2.937	2.969	2.957	2.929	2.984	3.022	3.017
[1248] China, Main, [S2514] MAIZE & PRODUCTS							
[5510] Production [t]	177,463,950	192,819,491	208,170,402	218,530,791	215,646,300	224,630,000	219,552,000
[5610] Import Quantity [t]	1,633,630	1,810,142	5,250,375	3,285,904	2,607,014	4,735,537	3,176,680
[5071] Stock Variation [t]	7,918,269	7,976,749	10,326,163	8,102,363	-37,733	1,716,991	-1,025,823

Figure 38: Save Back to the SWS