# **GGIG Graphical Interface Generator**

# **User Guide**

Wolfgang Britz, August 2010

- Version March 2021 -

The following user guide documents the outcome of a collaborative effort of University Bonn and the author. Larger parts of the Java code underlying GGIG had been developed over the years in the context of projects related to the <u>CAPRI</u> modelling system, which received considerably funds from the EU research framework programs. Following the general policy in CAPRI, the GGIG pre-compiled code can be used for other scientific projects as well without charge. The document comprises to a larger extent the content of earlier versions of the CAPRI user guide of which the GUI is now realized in GGIG.

The author would like to acknowledge the contribution of Alexander Gocht, vTI Braunschweig, to the CAPRI GUI coding efforts. All errors remain with the author.

# Content

GGIG Graphical Interface Generator	1
User Guide	1
Content	3
Overview	9
An overview on the GUI	11
Initialization: General interface settings	12
Necessary software	12
Entering settings	12
GAMS and R related settings	14
SVN related settings	15
Case one: Exploiter and runner	16
Usage for installation purposes	
Settings linked to the exploitation tools and the content of the include file	
Projects	
Concept of projects	
Switching between projects	
Generating a new project	
Editing an existing project	
Help system	
Starting GAMS from GGIG	
Viewing results: exploitation tools	
Views as the basic concept for exploitation	
Exploiting results	
The multi-dimensional viewer with pivoting and exporting possibilities	
Pre-defined views	

View selection	
Navigating through views	
Navigating in the outer dimensions of the viewport	
Column and row selection	
Predefined selection groups	
Selection of the view type	
Manually changing the pivot	
Changing view options: fonts, number formatting and rounding, hidin comparisons	ng empty cells, 
Showing a histogram window	
Working with tables	
The toolbar	
Tooltips for column and row headers	
Drill-down	
Clipboard export	
Export to file	
Working with EXCEL export	
Sorting	
Numerical filtering based on cell content	
Changing the row height and column width with the mouse	
Adding statistics	
Implemented outlier detection algorithms	
Pop-up menu	
Working with graphics	
General handling of graphs	
Setting colors manually	
"Walking" through the data	

Exporting the graphic to file59
Exporting the graphic to clipboard60
Pop-up menu in graphics60
Bar charts
Waterfall charts
Line, point and area charts64
Pie charts
Spider plots67
Box and Whisker charts
Histograms
Scatter plots
Deviation renderer
How to draw a line chart with mean / min / max etc. over a time series
Network graphs76
Markov charts77
Working with maps
Flow maps
Pie chart maps
Colored thematic maps
Changing the classification and the legend
Changing the coordinate set
Adding a histogram window to a map
Shrinking polygons according to UAA share
Area weighted classification
Excluding zeros from classification and removing small and large values
Classification method
Integration distribution information in the map window

Color table	
Changing the way the legend is drawn	
Changing the title of the map	
Zooming in and out and navigating in the map	
Getting data for specific polygons	96
Highlighting specific regions in the map	97
Updating the map	
Adding region label to the map	
Showing river and cities	
Storing and re-loading your settings	
Copying the map to the clipboard or to a file on disk	
Schemes shown with the mapping viewer	
Popup-menu in tables	
Exporting the data to ArcGIS	
Machine learning	
Motivation	
Implementation in GGIG	
Interaction between the GGIG GUI and WEKA	116
The WEKA GUI	
Classification	
Filtering	
Attribute viewing and selection	
Summary	
References	
Scenario editor	
Concept of the scenario editor	
Using the scenario editor	

Working with the policy editor
Tag editor
Search functionality
Update possibility for administrator
Meta data handling
Why meta data?131
Technical concept
File menu
Settings menu 135
Utilities and GUI menu
Utilities: Batch execution
Format of the batch execution steering file
Header
Settings for tasks
Program flow possibilities in the batch execution facility
Using the batch execution facility
The output from batch execution143
Automated GDXDIFF146
Error reporting and outputting results from GAMS to the HTML page
Utilities: Generate GAMS documentation in HTML pages
Structure of the HTML pages
Tagged in-line comments 150
Refactoring Consequences for Gams Code151
General overview
Example for a Symbol page
Example for a GamsSourceFile page
Example for a page for the a set 154

File list	154
Set element list	155
Utilities: Equation and variable viewer	155
Background and motivation	155
An overview on the viewer	156
Producing input for the view with GAMS	156
Includes	159
Loading symbols	160
Working with the equation and variable viewer	161
Decomposition	162
Utilities: Gdx-file(s) viewer	163
Utilities: Generating coordinate files for the exploitations tools from shapefiles	165
Analysis differences in GAMS based data using GGIG	168
Background	168
Comparing two data sets in GGIG, example from CAPRI	168
GGIG as GDXDIFF	169
Using the table definitions	172
Comparing two GDX files with GGIG	172
Index	172

## Overview

The GAMS Graphical Interface Generator (GGIG) is a tool to generate a basic Graphical User Interface (GUI) for a GAMS or R project<sup>1</sup> with five main functionalities:

- Generation of user operable graphical controls from XML based definitions. The XML file defines the project specific layout of the GUI. The user can then interact with the GUI to change the state of the controls. The state of each control component such as a checkbox can then be mapped to GAMS code (\$SETGLOBALS, Set definitions, settings for parameters). It combines hence the basic functionality of a GUI generator and a rudimentary GAMS code generator.
- 2. **Generation of GAMS compatible meta data** from the state of the control which can be stored in GAMS GDX format and later accessed, so that scenario definitions are automatically stored along with results.
- 3. Execution of a GAMS or R project while passing the state of the control to GAMS respectively R as a include file.
- 4. **Exploitation of results from GAMS runs** by providing an interface to define the necessary interfacing definitions in text file to load results from a GAMS into the CAPRI exploitation tools.
- 5. Access to a set of GAMS related utilities. This include e.g. a viewer for GDX files, a utility to build a HTML based documentation of the GAMS code or a batch execution utility.

That guide is thought for users of GGIG generated interfaces. It will be typically be complemented with a user guide which is specific for the project, such as the CAPRI user interface documentation. The "GGIG programming guide" comprises the necessary information to set up interfaces based on GGIG.

The main parts of GGIG are graphically depicted below. At its core stands the GGIG Control generator, based on Java code. Based on a XML based definition file provided by the project, it generates a project specific GUI which can be operated by the user. The state of these controls such as numerical settings, on/off settings or n of m selection can be passed to

<sup>&</sup>lt;sup>1</sup> The code can also be used from inside Java, but that feature is not discussed in the documentation.

GAMS by an automatically generated include file which also contains generated meta data documenting the state of the controls. The user can also execute GAMS from the GUI. The GUI can equally load numerical results and meta data in a specific GDX viewer. The latter supports "view definition", i.e. pre-defined reports to exploit the results. The details of the different elements are discussed below.



Diagram: Overview on information flow in GGIG

# An overview on the GUI

📓 CgeRegEU+ [t:\britz\capri\gams]					
File Settings Utilities					
CgeRegEU+ tasks	General settings Methodological switches				
🔘 Calibrate CGE	CgeRegEU+ General settings				
Run policy experiment					
Run test shocks					
Market States		Scenario	description cge_rd_plus10	•	
		Us	e seperate threads 🔽		
				_	
			AT000000 "Austria" TR000000 "Treland"	<u>^</u>	
			CZ000000 "Czech Republic"		
			SK000000 "Slovakia"	E	
		Countri	es DK000000 "Denmark"		
			HU000000 "Hungary"		
			FI000000 "Finland"		
			FR000000 "France"	*	
		Compile CAMS	Start GAMS Stop GAMS	Evoloit requite	
		Comple GAMS		Explore results	
1					
GGIG					
GAMS Graphical User Interface Generator					
UR Woitgang Britz 2012					
Institute for					
Food and					
Resource Economics	•	m			•
CgeRegEU+ Ini fil	e : regcge.ini User name	undefined User ty	oe : runner		

As shown in the example above, the GUI consists a few elements:

- 1. A **menu bar** which allows to change some settings (see the section on general interface settings)
- 2. A **workstep** and **task selection panel** on the left hand side where the user can select between different tasks belonging to the project.
- 3. A right hand side panel which either shows:
  - i. **The generated controls**, a button panel to start GAMS and a windows in which the message log from GAMS is shown
  - ii. A panel to select data to view and to start their exploitation
  - iii. The exploitation tools
  - iv. A scenario editor
- 4. A small window in the left lower corner which present a logo.

Before using a GGIG based interface, the users need to edit some project specific settings, see next chapter.

# Initialization: General interface settings

## Necessary software

GGIG is based on Java libraries. In order to use GGIG, a recent Java Run Time Engine, version 1.8 or newer must be installed and found on the "Path".

Typical reasons why the *GUI does not open* and possible remedies are described briefly below:

- *Java is not installed* or not registered on the PATH. In order to test that, open a DOS prompt and type Java. If you receive a message that the command is not found, (re-)install Java and make sure that Java also updates the PATH during the installation.
- A *wrong Java Runtime Environment* is installed. Please first check if Java version 8 is installed and not some older version (that check is not only useful for the GUI of GTAP8inGAMS, but also generally for security reasons if one uses Java plugins in web pages). *On 64 bit Windows operation systems, a 64 bit version of Java should be used.* You can test that again in a DOS prompt by typing "java –version".

#### java version "1.8.0\_25" Java(TM) SE Runtime Environment (build 1.8.0\_25-b18) Java HotSpot(TM) 64-Bit Server VM (build 25.25-b02, mixed mode)

The response should look similar as shown below (note the string "64-Bit"). Unfortunately, many 64 bit Windows machines use a 32 bit browser, and if you install Java from such as browser, the 32 bit version is downloaded. The link <u>JRE</u> points both to a 32 and 64 bit version, make sure you select the 64 bit version on a computer with a 64 bit OS. Installing a 64 bit version on a machine where the 32 bit version is already installed does no harm and makes the 64 bit version the default.

## Entering settings

The interface has a few standard settings which can also be accessed via the "edit settings dialogue". These are:

• **Certain file locations**: the directory where GDX files for results are assumed to be stored (resDir), and three directories which can be used to adjust the specific model application: the root of the GAMS file (workDir in GAMS ), called modelDir, a directory for restart files and one for data files. The application might not show all the

#### field or add additional ones.

🛃 Option		<b>— — X</b>
Option User Settings System Setting	35 GAMS and R SVN Other options	
model files directory	\gams	
Result Directory	T:\britz\promok2014\results	
Restart Directory	\restart	
Data Files Directory	\dat	
	Rename 'gamsName.lst' etc. by 'task.lst'	
	☑ Generate REF and EXP file	
Options	Save in caprinew.ini	

These file locations are passed to GAMS and can be used in the GAMS code to read / include files from the correct locations on disk. In order to make an initialization file portable, locations can be defined relative to the GUI directory.

The following screen shot shows entries in the include files generated based on the directory fields. The entries with a "R" at the end such as "resDirR" use unique-style forward slashes as delimiters and should be used with R-projects. The scratch directory stems from the GAMS settings tab discussed below.

#### **\$ONTEXT**

```
Directory settings from ini file
```

```
$0FFTEXT
```

```
$SETGLOBAL curdir '..\gams'
$SETGLOBAL curdirR '..//gams'
$SETGLOBAL resdir 'T:\britz\res_microdata'
$SETGLOBAL resdirR 'T://britz//res_microdata'
$SETGLOBAL datdir '..\dat'
$SETGLOBAL datdirR '..//dat'
$SETGLOBAL restartdir '..\restart'
$SETGLOBAL restartdirR '..//restart'
$SETGLOBAL scrdir 'd:\temp'
$SETGLOBAL scrdirR 'd://temp'
```

The "Look and Feel" (L&F), which can be changed via the menu bar:

\b	ritz\c	apri_liaise\gams]			
5	GUI	Settings Help			
rk	Ba	tch execution		ettings Modu	
la	Ge	nerate GUI geometry from shapefile		heral settings	
at	Lo	ok and Feel	►	System	
e	enario Meta			Metal	
ks	(S			Nimbus	
	cons	vio			

The system L&F is the default chosen and will render controls as close as possible to the standard of the operating system. The "Metal" and "Nimbus" are L&Fs which are portable across platform. Especially "Nimbus" has a nice look (compare with the screen shots above):

CAPRI [T:\britz\capri_liaise\gams]		and where	-	
File Utilities GUI Settings Help				
CAPRI worksteps	General settings Modules and algorithm Reporting Methodological switches CGE			
CAPRI worksteps  Build database Generate baseline Run scenario CAPRI tasks  CAPRI tasks  Define scenario Run scenario with market model Run scenario without market model Oownscale scenario results Run policy experiment with CGE	General settings       Modules and algorithm       Reporting       Methodological switches CGE         CAPRI General settings       Scenario description <ul> <li>scenTemplateGHGAbatement</li> <li>Generate GAMS child processes on different threads</li> <li>Use new global version</li> <li>EU28</li> <li>GHG abatement technology</li> <li>off</li> <li>off</li> <li>etchnology</li> <li>etchnology</li></ul>	) Vea Base year (20 Simulation yea	rs 004  2012  2013 2015 2020 2025 2030 2025 2040 2045 2040 2045 2050	Countries
	Comple GAMS Star	art GAMS Stop	GAMS ) Ex	ploit results
	GAMS output			

## GAMS and R related settings

S Option	the second se	_ 0 <b>_ X</b>
Option		
User Settings System Settings GAMS and R SVN Other op	tions	
Path to GAMS.exe	L:\24.3\gams.exe	
GAMS scratch Directory	d:\temp	
Path to R.exe		
Path to Troll.exe		
GAMS Options		
Number of processors used in GAMS	Get the number of processors	8
Processor speed relative (100% ~ 2.4 GH Intel core 2		100 👤
	Save in caprinew.ini	

The "Path to Gams.exe" points to the actual GAMS engine to use. It can be entered in the text field. Alternatively, you can use the button to the right of the field to navigate to the directory where "GAMS.EXE" is found via a file selection dialogue. Please do not only enter a directory, but the full file name as shown above (and choose GAMS.EXE, not the user interface of GAMS, GAMSIDE.EXE). Additional executables needed by the project might be shown additionally.

The "Scratch Directory" will be passed to GAMS and determines where GAMS stores temporary files. A directory on a local disk (not one on a file server) should be chosen.

The "GAMS options" field allows the user to send its own settings to GAMS, e.g. as shown above, the page width used in GAMS listings and the number of maximal process dirs generates by GAMs.

The number of processors used in GAMS will determine how many parallel GAMS processes will be started with threads are in use. The relative processor speed can be used by the GAMS code to e.g. determine if threads should be used or not.

The generated entries in the include file are shown below:

\$ONTEXT GAMS options \$OFFTEXT \$SETGLOBAL gamsPath L:\24.3\ \$SETGLOBAL gamsArg \$SETGLOBAL rExe \$SETGLOBAL rExe \$SETGLOBAL trollExe \$SETGLOBAL NoCPU 8 \$SETGLOBAL NoCPU 8 \$SETGLOBAL procSpeedRelative 100 \$SETGLOBAL JAVA ON

Note that depending on the application, other executables besides GAMS might be shown in the panel.

## SVN related settings

🛃 Option	
Option	
User Settings System Setting	gs GAMS and R SVN Other options
SVN user id	
SVN password	
SVN URL for Gams	https://svn1.agp.uni-bonn.de/svn/capri/trunk/gams
SVN URL for results	https://svn1.agp.uni-bonn.de/svn/capri/trunk/results/regcge
SVN URL for restart	
SVN URL for data	
	Report SVN URLs and local modifications to include file
	Save in caprinew.ini

The SVN settings can be used to perform checkout and updates in cases where the model code with related data, restart files or result files is under versioning control on a SVN server. If the model is not under version control, the settings "svn=no" renders the tabbed plan invisible. The SVN settings are thus only optional.

#### Case one: Exploiter and runner

#### Entering the necessary information to link to the SVN server

An exploiter by definition only accesses GDX files from the result directory. He is not allowed to run GAMS programs, and thus does not need access to the GAMS source code, data and restart files read in by the different GAMS based working steps of CAPRI.

Accordingly, in order to work with SVN, only three pieces of information have to be entered



under "Settings / Edit Settings"

- The SVN user id
- The SVN password
- The url of the result directory

The first two fields are not visible, and the related entries in the ini file are encrypted. The last entry can be set to a specific branch relating e.g. to a training session. That allows for CAPRI "mini installations". These mini installations do not need to be distributed as SVN installations as the SVN interface in the GUI will also allow to "checkout" over existing subdirectories and files. That ensures some additional safety regarding access information to sensible branches of the server – a bystander cannot read the user id and password. But users should always place local copies of such branches including the directory from which the GUI is started on secured parts of their file system.

The local directory for the GUI is simply taken from the start directory of the GUI; whereas the SVN address for the GUI is stored in the "default.ini" file.

ild database		Base	year	2004	
enerate ba	seline	Simulation year		2020	
ID COD DE				EU027	
JII SCENARIO				EU025 EU015	
ploit gdx (	🕌 Option				
	Option				
	User Settings CA	PRI System	Settings SVN Other options		
	SVN user id				
selection	SVN password		•••••		
ownscale s	SVN URL for res	ults	https://svn1.agp.uni-bonn.de/svn/old/trunk/results		
ploit scen			Save in capri.ini	]	
		Scen	ario 5		
				show results show meta data	

The runner can enter the additional SVN urls relating to the different sub-directories of a CAPRI installation. That should give some flexibility when working with branches on the server:

🙆 Option									
Option									
User Settings CAPRI Syst	User Settings CAPRI System Settings GAMS SVN Other options								
SVN user id	•••••								
SVN password	•••••								
SVN URL for Gams	https://svn1.agp.uni-bonn.de/svn/old/trunk/gams								
SVN URL for results	https://svn1.agp.uni-bonn.de/svn/old/trunk/results								
SVN URL for restart	https://svn1.agp.uni-bonn.de/svn/old/trunk/restart								
SVN URL for data	https://svn1.agp.uni-bonn.de/svn/old/trunk/dat								
	Save in T:\britz\CAPRI\GUI\capri.ini								

#### Performing an update

The second functionality for an exploiter (and runner) is to update all directories with the

	🚣 CAPRI [t:\bri	tz\capri\gams]	
	File Settings	Utilities GUI Help	
	CAPRI works	GDX Viewer	
	🔘 Build data	Start equation and variable viewer	
	Generate	Build HTML documentation	
	Run scena	SVN update	
ľ	CAPRI tasks	Use 't:\britz\capri\gams\util\cleanUpGamsDir.bat' to clean up source directory	. An update

```
menu item "Utilities / SVN update"
```

will download updated versions of files into hidden directories, and, if the related files in the local working copy have not been modified, will also replace the local files.

Choosing that menu item will open a dialogue with just one button termed "update" and an area into which messages from the SVN updates / checkouts are reported:

SVN settings	
	^
	≣
	<u> </u>
Update	

Pressing the "update" bottom will trigger an unpdate. Possible conflicts, merges etc. are shown in the reporting area:

Update for t:\britz\capri\gui completed at revision 5310		
Files/sub-directories updated : t:\britz\capri\results	Ι	
Update for t:\britz\capri\results completed at revision 5310		
Update for t:\britz\capri\dat completed at revision 5310 Skipped t:\britz\capri\restart\fert_FR.gdx		
Files/sub-directories updated : t:\britz\capri\restart\fert		
Skipped t:\britz\capri\restart\inputs\L&B_EXPTV&L.GDX		
Files/sub-directories updated : t:\britz\capri\restart\inputs		
Files/sub-directories updated : t:\britz\capri\restart		
2 files were skipped (probable conflicts)		
Update for t:\britz\capri\restart completed at revision 5310		
Update for t:\britz\capri\gams completed at revision 5310		
	Update	3

If the directory is not yet under version control, the GUI will perform a checkout instead, i.e. setting up the first installation of the hidden copies from the server. Before an update, a "clean-up" operation will remove any possible local locks related to earlier unsuccessful SVN operations. As long as an internet connection is available, that should ensure smooth updates in most cases and avoid some of the more tricky problems TortoiseSVN users might face.

#### Case two: Administrator

An administrator can enter the same SVN directories as a runner, but can trigger updates for the different parts separately:

ĺ		
	- SVN settings	
	ns.	<u>^</u>
		~
	Update GAMS Update results Update restart Update data Update GUI	

#### Usage for installation purposes

Since quite a while the CAPRI network discusses how installations specifically for training sessions can be organized more easily. The newly embedded SVN functionalities in the GUI

should ease that task somewhat, specifically in cases where only exploitation functionalities are asked for.

The installation of CAPRI based on the new functionality is relatively straightforward. As before, a JAVA run time engine must be installed for the GUI to run. For an exploiter, only a minimum GUI installation (e.g. without the large geometries for the 1x 1 km layer) and the necessary results files to view can then be copied to a local directory. At first start, the user must then only enter where the results had been copied to (if the result files are not parallel to the GUI) and save the information to his new *CAPRI.INI* file.

💰 CAPRI [\capri\gams]						<b>_ D</b> 🔀
File Options Version control He	lp					
-Work step selection	Set XML ta	ble definition file	Use table definitions from : null	Sort code lists	Show dialog to link GDX dimen	isions to sets
• Exploit gdx files	-List of tables loaded from GDX file	s)				
Exploit gdx files	List of tables loaded from GDX file	s) Nachric	tht Result directory '\caprily esults' does not a	X exist.	2	
			Load gdx file Loa	d selected tables(s		
CAPRI GUI Version 2.0 , March 2009	Ini file : capri.ini	User name :	User type : exploit	er		

The interface is set-up such that only the results of those work steps are visible where result files are found. For a training session concentrating on analysing scenarios, only those result files can be distributed. An installation with four scenarios at NUTS2 level plus all the necessary GUI files will require under 100 MByte disk space.

🛓 Option		
Option		
User Settings CAPRI Sy	stem Settings SVN Other options	
Result Directory	d:\TSCheltenham2010\results	
	Rename 'capmod.lst' etc. by 'task+settings.lst'	
	Save in capri.ini	

Once the user has optionally entered the results directory, and stored it to the ini file, the user will face a rather clean interface which only allows to exploit existing scenarios and to exploit GDX files (also that option could be removed for exploiters).

# Settings linked to the exploitation tools and the content of the include file

🔮 Option			— C	
User Settings FarmDyn System Settings	GAMS and R SVN Other options			
Path to view definition file (tables.xml)			Sort code lists in predefined tables	
Path to Editor	C:\kedit\KEDITW32.EXE		Sort code lists in GDX viewer and equation/var v	iewer
Title for interface			Clean window with GAMS output with each new GAMS of	compile/start
Language to load from tables.xml	English		Use task specific settings in interface	
✓ Output batch	n statements to include file		☑ Debug XML table definition (output to OS prom	npt)
Output ini s	statements to include file			
Output set with r	neta information to include file			
	s	ave in dairydyn.ini		

The "Path to the view definition (tables.xml)" allows loading a XML file which defines views into the results (see chapter "exploit results").

If you are using regularly a text editor, you can register it under "Path to Editor". That will allow to open the GAMS source file, listings and the generated include file in that editor directly from the GUI.

"Use task specific settings in interface" relates to how control settings are stored. If a control is present for several tasks (such as control to select the regions), the current setting can be either stored generally or specifically for the current task.

Normally, the order of the items in tables is defined by GAMS and/or the in the table definition file. The two sort options overwrite this behavior. Note that you can also search for items in the selection lists, which motivates why the default of these settings is off.

In order to facilitate the use of the batch facility, matching statements are normally outputted to the include file. To keep it shorter, this can be switched off. The same holds for the META information and the statement which can be outputted to the ini-file.

## **Projects**

## **Concept of projects**

In order to support the idea of a simulation environment which supports the application of a tool in a specific policy context, i.e. a narrative, the concept of a project was developed. A project combines a content related application context, i.e. a specific use case of the tool e.g. to a policy question or the development of a baseline, with the necessary technical information on that application. As generally with GGIG, that technical information is captured by the state of user accessible controls which might include file names. These files can then include large contents of information such as a release of data base.

A project in GGIG comprises three types of information:

- (1) Core information on the project itself such as its name
- (2) Project specific default settings for the GUI controls and
- (3) Project specific default settings for general properties accessible via the settings dialogue.

The user can switch between projects from the menu bar, doing so has three effects. Firstly, the project specific defaults for the control become active. If the GUI control was not yet operated by the user, the settings passed to GAMS, R or Java are the project specific defaults. Secondly, if the user operates a control, its input is stored project specific in the ini-file. If the GUI is restarted, it remembers the settings specific to the project chosen. Thirdly, context information on the project such as its webpage is added to the menu bar.

The concept of a project in GGIG hence firstly helps to tailor the GUI to a specific application context of the tool. The user can be informed about that application context (project name, start and end date, its web page, a PDF document, a contact point etc.). Equally, specific settings which should be used with the project (e.g. a specific version of the data base, a default shock file for the baseline, simulation years, a set of indicators) can be registered, such that the user does not need to introduce them by hand. Users who apply the tool in different application context can switch between projects from the GUI and automatically receive the context specific defaults plus potential updates done by them.

There is hence typically no need to install the tool several times when it is applied in different application contexts. However, it might be useful to have separate result directories for each application context. The concept of a project does fits seamlessly in the existing approach of GGIG and supports both end-users which might only exploit results and power user who often face the challenge to apply at the same time a tool in different application contexts.

#### Switching between projects

The user can switch between projects with the "Set project" menu item under the file menu

🕬 GTAP in GAMS V8 GGIG-GUI [\ga						
File	Utilities GUI Settings Help					
	Exit worksteps					
	Set project > New project					
	Edit project		None			
	New project					
	Font					

The currently active project is shown in the bottom info bar:



If a web page, a pdf document or an e-mail contact address is registered with the project, a new menu item is added to the top menu bar, from which information on the project can be retrieved:

```
      File Utilities GUI Settings
      Project Help

      GTAP in GAMS V8 works
      Webpage http://www.ilr.uni-bonn.de/agpo/agpo_d.htm..
```

Internally, any control settings are stored project specific, e.g.:

```
new\ project.import\ from\ gtapagg.relative\ tolerance=1.0
new\ project.import\ from\ gtapagg.reduced\ thresholds\ only\ in\ combination=false
new\ project.import\ from\ gtapagg.print\ gams\ code\ to\ listing=offListing
new\ project.import\ from\ gtapagg.filter\ method=Rebalancing
```

## Generating a new project

A new project can be generated from the menu bar with a menu items under the "File" menu in the top menu bar if the user has administrator rights:



Choosing the menu item "New project" will open the project dialog:

*** New project	$\times$
Core info Defaults System settings SVN	
Start date       Policy context       Pdf       Author       Name       End date       Mailto       Webpage	
Generate project	

The dialog comprises four tabs. The first tab "Core info" shown above comprises information on the project itself. The entries "Pdf", "Mailto" and "Webpage" generate new items in the top menu bar and provide information on the application context to the user.

The second tab "Defaults" comprises a list of GUI controls of the tool with their title and its default setting:

www. New project			$\times$
Core info Defaults System settings SVN			
Draw selection		mean	^
Regions with reduced thresholds			
Solprint		2	
Exceptions			
Options for solve outputs			
Symbol list with cross references		offSymxRef	
Penalty for deviation from		Private consumption, Government consumption, Inve	es
Number of stepwise increases of thresholds		10	
Symbol list		offSymList	
Minimum number of transactions		50000	
Print gams code to listing		offListing	
Relative tolerance .E		-2	
Options for listing			$\sim$
L	<u>G</u> enera	te project	

The left hand size shows the title of the control, the right hand comprises a text field with the default settings as stored in the XML-definition file for the GUI. That general default can be overwritten with a project specific one. Only these differences are stored-

Similarly, the two remaining tables show under "System settings" the location of system directories:

🕬 New project			×
Core info Defaults System	settings SVN		
Model files directory	\gams		
Result directory	\results		
Data files directory	\data		
		Generate project	

And, if SVN information is registered with the tool, the SVN fields:

🐡 New project		$\times$
Core info Defaults System	settings SVN	
SVN user id		
SVN password	•••••	
SVN URL for Gams	https://svn1.agp.uni-bonn.de/svn/Gtap8InGams/gams	
SVN URL for results		
SVN URL for restart		
SVN URL for data	https://svn1.agp.uni-bonn.de/svn/Gtap8InGams/data	
	Report SVN URLs and local modifications to include file	
	<u>G</u> enerate project	

When all edits are done, pressing the "Generate project" button will store the edited information in a separate ini-file which comprises the name of the project:

project\_New project\_default.ini

That ini-file comprises the information on the project and potential changes to default settings:

```
* * * Top of File * * *
#GGIG properties for project New project
#Sun Dec 13 09:54:04 CET 2015
New\ project.webpage=http\://www.ilr.uni-bonn.de/agpo/agpo_d.htm
New\ project.name=New project
* * * End of File * * *
```

That project file is loaded additionally to the normal ini-file when the user switches to the project. It can be found in the GUI folder and added e.g. to the SVN repository to make it available to other users. The GUI will scan the directories for such files and built from that list the projects available to the user.

That also implies that a user who simply wishes to handle several set of control settings via a tool can simply generate projects which only differ in name. Switching between those will than not affect default settings, but provide context specific storage and loading of all settings handled by the GUI (control settings, directories, selected scenarios for the exploitation tools, view settings).

## Editing an existing project

Once a project has been generated, a user with administrator rights can edit an existing project:



The dialog is identical to the one described above.

# Help system

Many user operable controls are linked to context sensitive help which can be queried with F1. That will open PDFs files.

# Starting GAMS from GGIG

GGIG allows starting the GAMS project directly from the interface, either in compile or run mode. A break request can also be sent to GAMS ("stop GAMS"):



Once started, the GAMS project routes its output instead to the console (the DOS prompt) to the lower right part of the interface, such that the user can follow the progress:

GTAP8 in GAMS : Calibrate parameters for 57x134 (11 se	ר(c)
comgms(5015) 700 Mb	^
comgms(5015) 725 Mb	
comgms(5015) 799 Mb	
comgms(5019) 820 Mb	
comgms(5019) 891 Mb	
com .gms(5019) 907 Mb	
comgms(5025) 938 Mb	
com .gms(5025) 939 Mb	
comgms(5029) 970 Mb	
comgms(5029) 963 Mb	
com .gms(5029) 1039 Mb	
comgms(5029) 1051 Mb	
comgms(5042) 1146 Mb	
comgms(5042) 1176 Mb	
com .gms(5047) 1211 Mb	
ccmgms(5047) 1212 Mb	
com .gms(5047) 1219 Mb	
ccm.gms(5047) 1220 Mb	
ccmqms(5047) 1225 Mb	
ccmqms(5047) 1226 Mb	
ccm	
com.(5047) 1233 Mb	
com .gms(5047) 1238 Mb	
com.(5047) 1238 Mb	
com.(5047) 1243 Mb	
	~
< > > > > > > > > > > > > > > > > > > >	

The left corner shows the name of the current tool (here GTAP8 in GAMS) and potentially a message sent from the GAMS process to GGIG (here "Calibrate parameters for 57x134"). The right hand side shows the time since the current process is started.

The pane with the content can be scrolled by a right mouse click in the pane to open a popup menu. If an editor is added under "other options", the GAMS and the listing file can be opened as well:

Open gams file .. Open gams Ist file .. Scroll Lock

The pane can hence be "frozen" so that e.g. the status of a model solve can be inspected while the project continues to run. During the frozen state, the messages from GAMS are not collected by the GUI, but outputted to the process window in the background instead. In order to successfully start a project, the ini file for GGIG must comprise the information where the GAMS executable can be found, but also where the GAMS code of the project to start is stored, see the discussion on settings above.

Note that if a break request is sent, it is probably not honoured (directly) by GAMS, rather, the program might continue to run until it finishes which might take quite some time. Therefore, the "Stop" button can be pressed a second time in which case, after 10 seconds, the pipe which collect the output from GAMS is closed and the GAMS program will typically immediately stop with an error message. In both cases, there is no guarantee that e.g.

temporary files generated by the currently running GAMS process are properly removed or closed.

## Viewing results: exploitation tools

The basic strategy of the GGIG exploitation tools roots in the <u>CAPRI exploitation tools</u>, which require that all model results are stored into one GAMS parameter which can have up to 10 dimensions and saved to GDX container as a sparse matrix on disk. One or several GDX containers with results are then read from disk and merged.

An additional dimension can be added if several files are loaded, e.g. to compare scenarios or years. A specific XML dialect defines views (filters, pivots, view types) into the cube, and allows the user to load several result sets – typically from different scenarios – in parallel.

If no table definition file is present, GIGG offers a GDX viewer which some interesting possibilities not found in the standard GDX viewer (such as numerical sorting, statistics, selections). For details, see below.

## Views as the basic concept for exploitation

The concept of the GGIG exploitation tools is centred on the idea of a view. Content wise, each view may be understood as showing one or several indicators relating to results of working steps defined in GGIG, e.g. environmental effects of farming, prices or market balances. Each view thus

- extracts a certain collection of numerical values (filtering)
- labels them so that they carry information to the user (long texts, units)
- chooses a matching presentation as a table, map or graphic
- and arranges them in a suitable way on screen.

The views can be linked to each other, allowing a WEB like navigation through the data cube. Views can be grouped to themes. The user may open several views in parallel, and she may change the views interactively according to her needs, e.g. switch from a map to a tabular presentation, or for instance change the pivot of the table, sort the rows, add statistics, introduce comparisons.

Internally, each view is stored in a XML schema. Technically, a view can be understood as a combination of a pre-defined selection query, along with reporting information. The XML

schema allows to attach long texts, units and tooltips to the items of a table, and thus to show meta-data information to the user. The XML schema does hence replace look up tables in a DBMS. It may equally store information regarding the pivoting, the view type (table, map, different graphic types), and for maps, classification, colour ramp and number of classes. The views can be grouped into logical entities, and are shown as a popup menu to the user.

Tabular views may feature column and row groups. Empty columns and rows can be hidden; tables can be sorted by column, with multiple sort columns supported. Numerical filter can be applied to columns.



## **Exploiting results**

For each work step, pressing the "Exploit results" button:

Compile GAMS	Start GAMS	Stop GAMS	Exploit results

will load the exploit result exploitation panel shown below.



Graph: The interface in "exploitation" mode

The right hand side comprises a set of drop-down boxes from which up to 15 different scenarios (or result files) can be selected. Each box comprises the list of GDX files found in the result directory registered for that task. The first element in each box is empty. The user can thus select in each box a file or leave it empty. The text fields next to the boxes allow the user to replace the file name normally used as a description of the scenario by a user chosen text.

The left hand side might shows, depending on the work step, for instance selection control for countries, the regional level, base year and simulation year. Operating these control filters out files from the disk shown in the drop-down boxes. In the example above, only results files for the base "04", simulation year "20" and the regional level 2 (= NUTS2) can be selected.

At the bottom of the panel, pressing the "show results" button will open the exploitation tools, the screen shot shown an example. The "Show meta" option is discussed in the section on "Meta data" and the GDX viewer under "Utilities".

Show n	neta	Show res	ults Loa	ad content of	f files into GD	X viewer	Return		
DAIRYDY	N [\gar	ns]							
View Handling	) Window	is							
🍝 Herd su	ımmary,	mean [0]							
TZEST Ye	ar		Farm					View I	:ype
	iean	~	RES 30COWS	2017				🕻 🔛 🖂 Та	ble
0		Cows			I	I			
T		Herd size [Heads]	Revenues [Euro/ha]	Variable costs (incl. concentrates)	Costs of concentrates [Euro/ha]	Gross margin [Euro/ha]	Labour [hours/ha]	Milk yield [liter/head and year]	Nu lac [lit
Ľ	( t			[Euro/naj					an
actBased	red0	25.71	2980.48	451.46	201.29	2529.02	22.00	7.14	_
actoused	red1	25.68	2973.75	449.12	199.42	2524.63	22.00	7.13	
	red2	25.01	2971.81	448.56	199.08	2523.25	22.00	7.13	
	red3	24.24	2975.55	449.85	200.11	2525.70	22.00	7.14	
	red4	23.44	2980.59	451.59	201.50	2528.99	22.00	7.15	
	red5	22.67	2985.35	453.17	202.68	2532.18	22.00	7.16	
	red6	22.01	2986.51	453.38	202.64	2533.13	22.00	7.16	_
	red7	21.35	2987.75	453.57	202.55	2534.17	22.00	7.15	
	red8	20.70	2989.09	453.73	202.35	2535.36	22.00	7.15	_
	red9	20.04	2990.52	453.89	202.15	2536.62	22.00	7.15	_
	red10	19.33	2992.08	454.18	202.11	2537.90	22.00	7.15	_
prodBas ed	red0	25.71	2940 22	437.50	190.10	2523.02	22.00	7.14	_
	red?	24.75	2938.59	437.07	189.89	2501.52	22.00	7.05	
	red3	24.10	2937.43	436.64	189.51	2500.79	22.00	7.04	
	red4	23.43	2938.74	437.02	189.74	2501.71	22.00	7.05	
	red5	22.80	2937.10	436.33	189.04	2500.77	22.00	7.04	
	red6	22.13	2939.59	437.03	189.44	2502.56	22.00	7.04	~
		<							>

The full functionality is only available if a table definition file (see programmer guide) matching the structure of the parameters in the GDX file is provided.

# The multi-dimensional viewer with pivoting and exporting possibilities

The results are – as mentioned above – loaded from one or several GDX containers and comprise the content of one or several GAMS parameters with up to 10 dimensions. The resulting data cube is loaded in a spreadsheet like viewer with pivot-possibilities.<sup>2</sup> The user may switch between a tabular view of the data, or different types of graphs (line, bar, pie, spider) or maps. Scroll-down boxes allow the user to rotate through data dimension not shown

<sup>&</sup>lt;sup>2</sup> Both the GDX container and the interface work in "sparse mode", i.e. only non-zero values require disk or memory space. Introducing additional dimension has therefore limited impact on space requirements. The viewer has proven to work with several 10 Million non-zeros.

in the view port columns and rows. Several data dimensions may be merged into one view port dimension. The user can use column and rows groups, and may apply selection to columns and rows as well as to columns and column groups. Rows and columns which carry only zero values may be hidden. Rows may be sorted by size of the numerical values in one or several columns. The content of the current table may be copied to clipboard. Alternatively, all or a selection of tables may be exported to an external file, in different formats (HTML, CSV, tab-separated, GAMS, fixed width tables). There are further possibilities available such as changing fonts or the number of decimals. The following chapters give details on the different functionalities available with the exploitation tools.

#### **Pre-defined views**

An XML file can define pre-defined views for the result content of the tasks defined in GGIG. Each view defines selections in the different data dimensions, the view type (table, graph or map) and the pivot, plus some other information. Pre-defined views thus can provide a guided approach to exploit model results, such as showing market balances, trade matrices, crop budgets, a welfare analysis or different environmental indicators.

#### Graph: A pre-defined view

🖄 Welfare overvier	w [0]		
	Region European Union	Year 2020 -	View type
		RES_2_0820ANIMWELF	
¥			
Total [Mio Euro]			20457780.00
Money metric [Mio Euro]			20435642.00
Agricultural inco [Mio Euro]	me		173051.80
Premiums [Mio Euro]			51902.55
EAA Output [Mio Euro]			495764.22
Output crops [Mio Euro]			229820.02
Output animals [Mio Euro]			265944.19
EAA Input [Mio Euro]			375461.16
Crop specific Inp [Mio Euro]	ut		96370.45
Animal specific In [Mio Euro]	nput		173872.92

#### View selection

The currently selected view is shown as a description of the window title:



The number behind the view gives the internal order of the views as several views can be opened in parallel.





## Navigating through views



The dark triangle to left and right of the view button allow navigating through to the list of available views. The outer triangles in grey allow navigating through the previously visited views.

#### Navigating in the outer dimensions of the viewport

In many views, some data dimensions will not be shown in the columns and rows, but as drop-down boxes in the toolbar. Use the mouse to select within the boxes. You can also use the keyboard to search items by typing. An example for these controls is shown here.

Region		Years
European Union 27	*	2013 🗸

#### Notes:

- If an outer dimension does only comprise one element, no drop-box list is shown
- If the toolbar gets too large, e.g. by having several drop-boxes with long item descriptions, such that its full content cannot be seen, it can be detached from the window by clicking on the line of small point under the icon and handled independently.

🛃 Supply	details, mapping view [0]	
1		
😤 Supply details, mapping	g view [0]	
	Scenario : RES_2_0820MTR_RD	
	Toolbar Supply details, mapping view (0)	

• You can also change the layout of the toobar which is especially helpful if there are many selection boxes for dimensions:

Options	
Mi	nimal toolbar on/off
Se	cond toolbar on/off
Se	cond toolbar in separate line on/off

The impact of these options is shown below:

Default toolbar



Tool bar minimal (i.e. some fast access button are removed)

🛃 Demand, by product [0]				
Regions	Items	Origins	Years	Percentage diff. to
World	V Quantity V	Total 🗸	shock	noShock

Second toolbar, shown in second line (selection boxes in separate line)

🛃 Demand, by product [0]			- • •
Percentage diff. to Scenarios noShock		- 🗁 🔧 🕻	View type Table
Regions	Items	Origins	Years
World ~	Quantity	✓ Total ✓	shock

#### Column and row selection

The user can selected select specific columns and rows (groups) to be hidden or included in the current view by using the buttons with the filter symbol shown below. Alternatively, selection buttons in the option dialogue can be used, see "Changing view options: fonts, number formatting and rounding, hiding empty cells, comparisons" below.



Double-clicking on one of these buttons will open a selection dialogue for the elements in that view dimension:



The selections can be done by mouse, following the convention of the operation systems. Additionally, a selection string can be entered in the field above, with the following possibilities:

- "\*" select all
- "C\*" select all items starting with "C", "C???" will select a string starting with C followed by any 3 characters.

After entering the selection string in the text field, one of the three buttons must be right-

clicked. The	Clear selection, add pattern to labels	button will remove	any selection and select
only those ite	ms which match the pattern enter	red in the text field.	Add pattern to labels will

keep the selection and add the matching items, whereas Remove pattern from labels will remove matching items from the selection.

Notes:

- 1. The selection dialogue can also be reached by view options dialogue, see below.
- 2. Alternatively, the pop-up menu of the view (right mouse click in the table, graphic or map) can be used to open the selection dialogs:

0.74	
0.92	
0.92	
0.92	
0.04	1
Reload	
Copy to Clipboard	
Export Data	
Pivoting	
Back to Table View	
Customize	
Selection 🕨	Columns
View P	Column groups
Statistics	Rows Row groups
Classify	
Table View	
0.93	
0.93	

#### **Predefined selection groups**

For some tables, pre-defined selection groups are stored. When the mouse is moved over the selection button and rests there for some time and such groups exists, a dialogue will show as below where the groups can be selected.

🛎 CAPRI [t:\britz\capri\gams]						
Π	View Handling Windows					
	🛓 Product Balances [	0]				
I	Region	Years				
	Germany	202	ð			
I	0	Selection of predefined groups				
		List of key words for groups				
I	₹	show all	^			
I	Cereals Double click	EAA (115) NITF (1)			p in	
	Oilseeds	NUTBAL (3)				
	Other arable field crops	SUGB (1)			3	
	Vegetables and Permar	all (145)				
	All other crops	dairy products (11)				
	Fodder	feed (12)			26	
	Meat	market (63)				
	Other Animal products	meat (4)	*	Cancel	3	
1	Voung animalo				5	

## Selection of the view type

As discussed below, the data can be shown as tables, graphics or maps, to do so use the view selection drop-down box:
		view type	
1		Table	~
		Table	<b>X</b>
		Line chart	, in the second s
	Crop observation	Bar chart	
	EV or 0.01 anir	Area chart	
	1.0 01 0101 0101	Spider chart	
		Pie chart	
		Box and Whisker chart	
		Histogramm	
		Мар	
		Flowmap	
		Markov chart	
		3.67	

## Manually changing the pivot

Normally, the predefined views will link the data dimensions in an appropriate way to columns and rows. However, the user is free to change the pivot, to e.g. generate a cross-

sectional series. A dialog opens when *right-clicking* the **button** to pivot the currently shown or selected part of the view (not available with maps):

🚳 Transposing and Merging	
Table control area	
Animation	0
(Box (1)	273
P - i - (070)	
Region(273)	
(Box (2)	1
hide(1)	
(Box (3)	1
Years(1)	
Box (4)	0
5511(1)	
Table area	
Table area	
2 Table column grou	ips
Scenarios(2)	
• Table columns	Item(6)
0 83	•
Table row groups Table rows Table cells area	
Activity(83)	
	]
ak	cancel
04	curror

The boxes **Scenarios(2)** show the data dimension and their lengths. They can be dragged to the different viewport dimensions as shown in the screen shot above. Assigning several

		Table columns			
dimensions	to the columns	Scenarios(2)	Item(6)		leads to "spanned"
unnensions	to the columns				leads to spanned
	BIOF_D2E2 Income				
	[Euro/ha or head]				
dimensions					

Alternatively, columns and rows can have additional block:

		•	Cereals	Oilseeds	Ot cro
BIOF_D2E2	Income [Euro/ha or he	ad]	438.65 -17.73%	454.55 -30.01%	
	Hectares or herd size [1000 ha or hds]		59217.85 -2.08%	8335.37 -13.24%	
	Yield [kg or 1/1000]	head/ha or head]	4958.80 -3.58%	2271.59 -6.17%	
BIOF_D10E1 0	Income [Euro/ha or he	ad]	533.21 0.00%	1 649.44 6 0.00%	
	Hectares or herd size [1000 ha or hds]		60473.37 0.00%	9607.46 0.00%	
	Yield [kg or 1/1000	head/ha or head]	<b>5142.97</b> 0.00%	2421.03 0.00%	

In combination with the selections for columns and rows, and column and row blocks, each view can be adjusted to the need of the user, e.g. to export the data in a specific ordering to an external file.

The pivot can alternatively be changed by mouse clicks in the text field above a selection box:

- A left mouse click puts the items from the selection box into the rows, while moving the items from the rows into a selection box. A double left click generates row groups.
- A right mouse click puts the items from the selection box into the columns, while moving the items from the rows into a selection box. A double right click generates column groups.

The row and columns can be switched by a left click on the "pivot button".

Alternatively, the pivot dialogue can be reached by the context pop-up menu of the view.

# Changing view options: fonts, number formatting and rounding, hiding empty cells, comparisons

A dialog opens when pressing the button to change various options of the current view:

S Customize view	
Tahoma	V 12 V plain V
Fraction digits and decimal	separator 2 V
Separater between merge	
Separator between merge	
Selection for : Activity	Selection for : Item
Column width 108 🚔	Row width 109 💼
Hide empty rows	Hide empty columns
Cut off limit to determine er	npty cells 0
✓ Use default pivoting for tables	istogram Use classification colors for tables
Show only selected items	Long texts only
Comparison output	Only values
Comparison threshold to hide values	•
	Region
Data dimensions used for comparisons	Item
	Year Scenario
Element used for comparisons Bagion	Dearmark
Element used for companyons Region	
ok define colors define statisti	cs store settings load settings

- *Fonts*: set font family, size and style affects tabular views, but also the different dialogues.
- *Number formatting*: chose the number of digits and define the decimal separator. The tool supports rounding numbers before the decimal point by allowing for negative fraction digits. Choosing e.g. -1 will round all numbers to tens. The numbers shown in graphics or tables are based on the rounded results is applied.
- Selections for columns, column groups (if present), rows and row groups (if *present*). The buttons will open selection dialogues, see the section on "Column and row selection" above.
- *Manually settings for column and row width*. The buttons will open selection dialogues, see the section on "Column and row selection" above. Alternatively, the row and column size can be changing by dragging with the mouse, a double arrow

resize cursor indicates that mode. If the column size is dragged while the shift key is down, the new column size will be applied to all columns.

- *Hide empty rows and hide empty columns* will suppress in the currently seen view, any columns and rows which would show only blank cells.
- *Cut off limit to determine empty cells* . In standard mode, the interface will treat zeros as missing values, and items will be shown as blanks. But the user might also enter a

different value (any value, in absolute terms, below the threshold set by the user will be treated as if it was zero).

- *Use default pivoting for table*: That is the normal mode, where the pivot is defined by the table views. By clicking that off, the currently chosen pivot (from the current table or manually defined) will be kept even if a different table is chosen.
- *Show histogram*: A histogram is shown additionally to the current view as a separate window. The current window might however hide the histogram window. The "Windos" menu can be used to bring the histogram windows to the front.



- *Use classification colors for tables* : Use the colours which would be used to colour the regions in a thematic map to colour the numbers shown in tables.
- Use of short code and/or long texts : Normally, long texts are shown to the user. Experience model developers might prefer to use with the actual GAMS code (short code).
- *Comparison output*: the exploitation tools can add different types of comparison output. They also affect what is shown in maps and graphics.



"Normalisation" means that the value is divided by the comparison points, allowing e.g. also to calculate shares. The "GTAP" difference is a compromise between a percentage and an absolute difference: it multiplies the difference in the logs with the difference (thanks to Rob McDougall from the GTAP team in Purdue for the proposal).

In tables, the "and" options will show two lines in each data cell, one with the observations, and one with the comparison output, as seen below.

		BIOF_D2E2							
Ø		Income [Euro/ha or	Hectares or herd size	Yield [kg or 1/1000	Supply [1000 t or 1000				
	•	head]	[1000 ha or hds]	head/ha or head]	animals]				
Cereals		438.65	59217.85	4958.80	293649.34				
		-17.73%	-2.08%	-3.58%	-5.58%				
Oilseed	s	454.55	8335.37	2271.59	18934.54				
		-30.01%	-13.24%	-6.17%	-18.60%				
Other arable crops		1693.24	7938.25	32743.16	259923.47				
		-4.26%	2.27%	-2.99%	-0.78%				

The "Data dimensions used for comparisons" allows to select one or several dimensions used for comparisons. A typical application is the comparison against a benchmark scenario. The use of more than one dimension allows e.g. to compare values against a specific year of a benchmark scenario. For each comparison dimension, a drop-down list allows to select the "Element used for comparisons" defined the comparison point. If statistics had been added to a view, these can be used for comparisons as well, they can be found at the bottom of the selection list.

#### Showing a histogram window

The system offers different ways to retrieve information about distributions of shown data. For maps and tables, the user can show an additional window with a box and whisker diagram, a histogram and some descriptive statistics as shown below. The box and whisker diagram is defined as follows: the green box shows the first (Q1) to third quartile (Q3), so that the width of the box is equal to the so-called inner quartile range (IQR). The blue "whiskers" are defined by Q1 minus 1.5 times IQR and Q3 plus 1.5 times IQR, but bounded by the minimum and maximum of the observations. In many applications, any value falling outside that range is classified as a mild outlier. The red dotted whiskers are at Q1 minus 3 times IQR and Q3 plus 3 times IQR, but bounded by the minimum and maximum of the observations. In many applications, any value falling outside that range is classified as a stronger outlier.

The user can restrict the plotted range as to exclude stronger outliers. If outliers are present, the red dotted whiskers at the tail with strong outliers are removed.

The blue dotted lines show the mean, and +/- one standard deviation around the mean. For a normal distribution, that would cover around 2/3 of the observations. The black dotted lines in the histogram show the class limits used for the colour model. The bottom reports some descriptive statistics. The technical implementation is set up according to the way maps are drawn: the population consists of all values in the rows and the columns of the table, and thus differs from the outlier control, which treats each column as a separate set of observations.



The colours are typically used to visualize the distribution in maps, but, as a second option, they can also be applied to the numerical values in tables. Alternatively, histograms and box and whisker diagrams can be drawn via the graphics.





# The toolbar

Region	Years	Scenarios		Ba III.	3	View type
Germany	2020	01GWP	*		~	Table

Note: with a minimal toolbar layout, the buttons won't be shown.

# Tooltips for column and row headers

For predefined tables, tooltips may be stored which give additional information on the columns and rows. They will appear when the mouse is moved over the respective column or row header.

	Y				
	Income [Euro/ha or head]	Hectares or herd size [1000 ha or hds]	Yield [kg or 1/1000 head/ha or head]	Supply [1000 t or 1000 animals]	Crop share [% or 0.01 a
_	Premiums + Reve	1484.26 nues - variable costs according to the definition of	6228.03 F Economic Accounts for Agriculture, income available	able to farmers to pay for land, labour, capital an	d profits
	550.57	65.49	2943.70	107.02	
_	1171.26	140.57	39192.45	5509.33	
	30713.87	14.42	285446.28	4116.05	

# Drill-down

Some views comprise hyper-links to other tables. Numbers with hyperlinks are shown in blue

2545.00	
2010.60	
22 De	Able click to table : Supply details

, and a tooltip will appear when the mouse is moved over them. Double-clicking in the cell will jump to the connected table.

# **Clipboard export**

The content of the currently shown view can be copied to the clipboard by pressing the button. Tables are placed as tab delimited text in the clipboard, so that they can be pasted into spreadsheets. Graphics and maps are placed as graphics in the clipboard and can be copied e.g. into word processing.

**Note:** If copying numbers from the clipboard to EXCEL, it might be necessary to change the delimiter to reflect the country specific formatting of numerical values.

If the clipboard content is pasted into a spreadsheet program, e.g. to generate nicely formatted tables or graphics, it is recommended to past the raw data into one sheet from clipboard and place the formatted table or graphic in another one, and use references to the raw data copied from the interface. If an update of the raw data is then necessary, for instance after the simulation was repeated, the old data might then be simply overwritten with a new paste from the clipboard and the formatting information is not lost.

# **Export to file**

A dialog opens when pressing the **button** to export the full dataset of the view to a file – not only the currently seen part.

The action provoked by the button depends on the view type. In tabular view, in opposite to the clipboard export, the export file will scroll through the outer dimensions and will copy all stacked tables after each other into a file. Take the table below as an example. Clipboard export will export the data for Belgium and 1984. File export fill export data for *all* regions and for *all* years, if the user does not apply filters in the export dialog.

-					
ĺ	🕌 Prepare natio	nal database			
	View Economic Account	For Agriculture	Region Years Belgium Viars	Table -	
	• •		RAW		
		•	Unit value EAA [Euro/t]	Quantity [1000 t]	
I	Production value				
I	Cereals			2515.60	
I	Oilseeds			22.00	
I	Inputs				
I	Premiums				
I	Production value				
I	Cereals			2515.60	
1				1337.40	

Generally, the data underlying one or several views can be exported to a wider range of format as shown below. However, the user should be aware that not only the currently visible selection will be exported, but as a default, all tables which can be generated by choosing any combination of elements shown in the drop-down boxes. That can lead to large numbers of tables being exported. If only the content of the current table is to be exported, a clipboard export is generally recommended. Alternatively, the user can use selections (see below), but these only affect the export of the current table, and no table should be selected as discussed below. In order to export the data in a map or a graphic, e.g. to GIS system, the view must first be changed to tables. Afterwards, the 🖬 button will open a file dialog as shown below. For GIS-export, e.g. to ArcGIS, DBF is the recommended format.

1	Please choose a file format for export					
E B						
	Select Type of Export					
	Select the format of the file to export					
	HTML (basic table format for internet browser))					
	(DBASE data base file, e.g. for ArcGIS or EXCEL)					
	(comma separated text file, for e.g. EXCEL)					
	🔘 (as GAMS table)					
	TABF (text format, tabulator separated)					
	TXT (text format, fixed field width)					
	Bin (binary format, for exploitation with exploit applet)					
	MSAccess MS Access Database					
	XLS (MS Excel workbook, 2003 and before)					
	XLSX (MS Excel workbook, 2010)					
	Next					

Once next is pressed, the next pane will open a file dialog to choose a file. In the case of export to a Microsoft Access Data Base, the file must exist.



Wolfgang Britz, Version March 2021

If desired, the **Next** pane allows openening selection lists for the different data dimensions. These selection only work if the only the currently active table is exported, in

which case the table selection (next panel) should be left empty.

🛓 Please choose a file	e format for export	-	×
Export Data			
Set Export Dimensions	5		
	Export selection for Region	Select	
	Export selection for Items	Select	
	Export selection for Origins	Select	
	Selection will only work if only cur	rent table is exported	
			Back Next
			HEAL
	17 / 28 / 17 / 17 / 17 / 17 / 17 / 17 / 17 / 1	011 7	1 30 40 40

You can

Select

next the tables for export.

	Please choose a file format for export
	Export Data
	Set Tables to Export
	Leave selection empty to only export current table and to use selections
	Markets : Demand, by institution
	Markets : Demand, by product
Ы	Markets : Demand, by product, bar charts for scenario comparison
	Markets : Demand, by product, per capita in \$
	Markets : Demand, by product, per capita in \$, map
	Markets : Factor demand
	Markets : Intermediate demand
	Markets : Markets, balance
	Model overview : Flow schemer
	Model overview : Income
	Model overview : Income, per capita
	Model overview : Income, per capita, map
	Model overview : Model properties
	Model overview : Price schemer
	Model overview : Regional aggregation
	Only export aggregates for products and activities Only export aggregates for regions
	Back

**Beware:** the pre-defined table structure will be lost with DBF format, as will the long-texts and units attached to the tables. However, in the case of DBF-export, a second file with that

information will be automatically created. If you solely want to export the table you have currently up front, use the "copy to clipboard" 🖹 button. The clipboard export will retain the pivoting and further information.

쵩 Please choose a file format for export	×
Export Data	
Start Export	
Maximum number of non-zero items to export :	6961227
Open File in Editor after file was created	
Define column	
List output - no data dimension in columns	
List output - no data dimension in columns	
Regions and HSMUs	
Activities	
Input and outputs	
Scenario	
· · · · · · · · · · · · · · · · · · ·	
	Back Start

The last pane let you decide for DBF-export if you want a list, or if you want the data dimension spanned across the columns. For exporting the HSMU tables, it is recommended to put "Inputs and outputs" in the columns.

In case of DBF format, if everything has worked well, we should now find two files: one with the data, named as chosen in the file dialog, and a second one with "\_meta" introduced before the file extension.

# Working with EXCEL export

If the format XLS or XSLX is selected with the data export utility, each view (table) selected by the user will be exported to a separate sheet. A subsequent export to an existing workbook will overwrite only these sheets which match the names of currently exported views. The sheets will take over formatting options from the exploitation tools such as the number of digits, or if percentages or absolute number are selected. As sheets are overwritten, the user should store e.g. graphics or formatted tables on other sheets and use references.

	А	В	С	D
1	Demand, by institution			
2				
3		Table x: Demand, by institution		
		: " Region : 'World'		
		Items : 'Quantity'		
4		Origins : 'Total' : "	'Total'	'Grains and C
5				
6				
7		'Total'	######	1
8		'Household demand'	######	
9		'Government demand'	######	
10		'Intermediate demand'	######	1
11		'Investment demand'	######	
12		'Export demand'	######	
13		'International transport demand'	542,00	
14				
15		Source : GGIG, 13. Dezember 2014		
16				
17				
18				
19				
20		Table x: Demand, by institution		
		."		
<b>1</b>	🗆 🖻 Demand, by institu	ition / Demand, by product / 况 /		

# Sorting

The rows can be sorted by one or several columns by clicking with the left mouse button in the column headers. Repeatedly clicking will change the sort order and then return to the unsorted list.

Adding additional sorting columns is achieved by pressing the "shift"-key and then using the mouse as explained before. A sorting symbol will show sort direction, and its size will show the sorting order.



# Numerical filtering based on cell content

Clicking with the right mouse button on one of the column headers will open the "filter dialog" which can be used to apply numerical filters to remove rows not matching the filter from the view.

🛎 Filter dialog	
Define numerical selection filter fortable rows Comparison operator	Comparison value
Clear selection and select according to filter	
Add result of filter to existing selection	
Remove result of filter from existing selection	]
OK Cancel	

### Changing the row height and column width with the mouse

While dragging with the mouse the bottom of the first row header, the cell height of each row (the height of each row) is changed at the same time. But, the column width can be changed selectively per each desired column (if you change the width on one column, the widths of the other columns do not change). The column width can be changed in a similar way by dragging the right border of the column header. Alternatively, the column and row width can be set in the "Changing view options" dialogue.

3	& Bui	d regi	onal database		💰 Build r	egion	al database			💰 Bui	ld regi	onal database			💰 Bu	ild regi	onal database		
v (	Prod	uct Bala	Region Danmark	×	View Product	Balance	Region s Danmark	×	]	View Proc	luct Bak	Region nces Danmark	• •	]	View Pro	duct Bala	Region nces Renmark	Table 🗸	
ſ	0	8 >	Y		0	>	Y			0	•	Y			٥	= >	Y		
ſ	<b>B</b>	8 >	Supply [1000 t]			>	Supply [1000 t]				•	Supply [1000 t]	ŀ	<b>→</b>		•	[1000 t]		
6	Cereal	3		9244.00 🔨	Cereals			9244.00	^	Cereal	ls		9244.00	^	Cerea	ls	9244.00		^
				-	Oilseeds	t		250.54	۲	Oilsee	ds		250.54	E	Oilsee	ds	250.54		
					Other and	*		4371.75		Other	arable		4371.75		Other	arable	4371.75		
•	Dilsee	is		250.54	field crons					Veget	ables		248.46		Veget	ables	248.46		
					Vegetable	8		248.46		All oth	er		5005.67		All oth	ner	5005.67		
				4274 75	All other			5005.67		Fodde	r		38944.26		Fodde	r	38944.26		
f	ield ci	ops			Fodder	_		38944.26		Meat			2168.92		Meat		2168.92		
					Mont	_		2168.92		Other	Animal		7655.62		Other	Animal	7655.62		
Ŀ	Innet	hlee		248.46	weat					Young			25601.70		Young	1	25601.70		
	ind	bica.			Other Anir products	nal		7655.62		Manur	e		1116.83		Manu	re	1116.83		
ľ	°erma crops	ient		~	Young			25601.70	~	Fertilis	ser		965.77	~	Fertili	ser	965.77		~

# **Adding statistics**

The user may add different statistics as rows to the table as reported in the following table. The observations are assumed to be mapped into the rows of the current views. Zeros can be treated as missing values. The statistics summarize the observations (=rows) separately for each column.

**Note:** The statistics are only calculated for the cells currently visible. Trying to show e.g. percentage differences of statistics against items not visible (e.g. a region in drop down box which is not the currently selected one) will not yield usable results.

Statistics	Shortcut
Sum over observations	Sum
Number of observations	nObs
Arithmetic mean	Mean
Median	Median
Standard Deviation	StdDev
End value in first quartile	q1
First value in fourth quartile	q4
Minimum of the values	min
Maximum of the values	max
Minimum limit for outlier detection as	minOutlier
defined from user settings	
Maximum limit for outlier detection as	maxOutlier
defined from user settings	
Free chosen algebraic expression	freeEval

The above listed statistic options can be either found by pressing the "define statistics"

statistics button in the "customize dialogue" box, which opens by clicking the button on the toolbar, using the "define statistics" button, or by right clicking on any cell inside the table to open the popup menu, and choosing "Statistics".

📓 Customize view	3
Arist       Il       plan       Il       Plan       Il       Il       Plan       Il	
Hide empty rows Hide empty columns	
Cut off limit to determine empty cells 0 🗢	Reload 3450.1
Use default pivoting for tables	Copy to Clipboard Export Data 3450.1
Characteristic and the second se	Pivoting 794.1
Snow only selected items Long texts only	Statistics 8706.1
Comparison output Only values	View View
Data dimension used for comparisons Region	Table View
Element used for comparisons European Union 15	6088.1
ok         define colors         define statistics         store settings         load settings	6049.1

The dialog allows selecting the different statistics as shown above and a outlier detection method as described in low. The settings related to the outlier detection method are dynamically changing depending on the chosen detection algorithm.

🛃 Set statistics	x
Image: Treat zeros as missing values for statistics       Only show outliers         Set maximum percentage of outliers       2 -	
Set outlier detection method No outlier detection	-
Select statistics Sum Nobs Mean Median StdDev q1 q3 min max minOutler maxOutlier freeEval	
Free evaluation field           ok         Cancel         Update	

The free evaluation field follows the convention used for the definition of <eval> fields in table definition, see programming guide, that option targets experienced users and can be used to calculate an additional item in each column on the fly derived from the results to be added as an additional row. A possible application is to e.g. normalize a specific item and then use the options to calculate percentage difference against the normalized results. It should be noted that the evaluation field has to define items for the dimension(s) mapped into the rows to avoid arbitrary results.

The selected statistic will appear as the first rows of the table:

🛃 Supr	oly details, mapping view	F01	
	Production activity	Item	
۲	Cereals	Income [Euro/ha or head]	
			01CWDBIID
		<b>*</b> ▶	UTOWPDOK
Nobs	÷		301.00
Mean			573.32
Median			550.26
median			625.66
StdDev			440.20
q1			410.20
q3			687.49
min			-533.03
max			1667.07
minOut	lier		-678.00
maxOut	lier		1824.64
Europea	an Union 27		646.43
Europea	an Union 25		672.55
Europea	an Union 15		732.82
Europea	an Union 10		534.48
Relaium			824.52

Perhaps the most interesting option is to show only the outlier rows besides the statistics in the table, as illustrated below:

🕌 Build regional database [Data	a View 1]		X
Table	Production activity	Years View type	
Supply details, mapping view	Other arable crops	Y Y Table	*
0	5	Income [Euro/ha or head]	
Nobs		270.00	
Mean		3164.16	
StdDev		40975.38	
Utrecht		488074.25	
Acores		109324.08	
Ahvenanmaa/Aaland		-438378.94	
		_	

# Implemented outlier detection algorithms

The GUI offers different outlier detection algorithms as discussed below. For all the methods, the user may additionally define a maximum percentage of observations shown in which case only the largest or smallest outliers according to the outlier detection algorithm selected will be added to the table view.

### Standard deviation around the mean

The user can define the factor  $\beta$  applied to the standard deviation. Observations are marked as outliers when their distance to the arithmetic mean exceeds the value defined by the multiplication of the standard deviation  $\sigma$  and that user defined factor  $\beta$ :

$$\neg \Big(\bar{x} + \beta_{\sigma} \cdot \sigma > x_i > \bar{x} - \beta_{\sigma} \cdot \sigma\Big).$$

That detection works well if the data are (approximately) normal distributed, but might fail if large outliers are present which can easily bias the results as they will change both the mean and the standard deviation of the observation sample. Further on, many time series e.g. in the CAPRI data base have by definition a lower limit of zero, so that the assumption of normally distributed data sets cannot hold. Therefore, other outlier detection methods are also implemented as discussed below. The dialog allows changing the factor  $\beta$  from its default of 2 which covers 95% of the values for normally distributed data.

#### Standard deviation of values normalized by median

The values are all divided by the median and the new series is classified as under the option discussed above. The main advantage of that method is the shift to a mid point which is less vulnerable to large outliers in the observations.

#### Standard deviation of trend line error

A regression is estimated by using the index position in the unsorted values as explanatory values. The resulting errors are then classified according to the first option discussed above. The typical application would be a table where consecutive time points - e.g. years - are shown along the rows.

#### Median and inner quartile range

Box-and-whisker charts, which are also supported by the graphics view, are using the median and quartiles to visualize the distribution. They are also an easy and robust way to detect possible outliers. First, the so-called "inner quartile range (IQR)" is calculated as the difference in values between the beginning value of the first and the ending value of the third quartile. The IQR then consists of the 50% range of values around the median. The IQR is often used as a robust replacement of the standard deviation, similar to using the median as a robust estimator for the mean. The IQR is multiplied with a user defined factor  $\beta$  added to Q3 respectively subtracted from Q1 to define the lower and upper bound for regular values. The default value for factor  $\beta$  is 1.5. In opposite to the mean and the standard deviation, the quartiles and the median are not affected by outliers at the tails of the distribution, allowing for a rather robust way to filter outliers:

 $\neg \left(Q3 + \beta_{IQR} \cdot IQR > x_i > Q1 - \beta_{IQR} \cdot IQR\right)$ 

#### Conformity based on relation of distances

Here, the following formulae are used, taken from Last & Kandel (2001):

$$\mu_{i,l} = 2 / \left( 1 + \exp\left(\frac{\beta \cdot n \cdot m \cdot (x_{i+1} - x_i)}{x_{i+m} - x_{i+1}}\right) \right)$$
$$\mu_{i,h} = 2 / \left( 1 + \exp\left(\frac{\beta \cdot n \cdot m \cdot (x_i - x_{i-1})}{x_{i-1} - x_{i-m-1}}\right) \right)$$

They define "conformity" from below and above by comparing the distance from the current value to its neighbour in relation to the average distance for a predefined group size *m*. Before the formulae are applied, the values are sorted. In opposite to the outliers based on first and second moment, the method is also able to detect outliers in between "clusters" of values. Inside such a cluster, differences in distances between values are small, so that the relation between the distance to the next neighbour, and the average distance between the neighbour and its m-th neighbour is around unity. The big advantage of the approach is that it does neither assume a certain functional form for the distribution (as in the case of the mean/standard deviation approach), nor a uni-modal distribution as in the case of the IQR method, and it is rather easy to compute. It may be worth to continue with a literature research in the direction of similar outlier detection methods.

The factor  $\beta$  describes how distances between succeeding values are assessed. Outliers are defined when the maximum of the above and below conformity is above a predefined threshold  $\alpha$ .

 $\neg (\max(u_{i,l}, u_{i,h}) > \alpha)$ 

Last & Kandel have tested their algorithm for  $\beta$ =0.001,  $\alpha$ =0.05 and m=10. There seems to be a rich literature on that kind of "neighbourhood distance", where outlier control based with different algorithms is analyzed in detail. The different parameter can be set by the user interface.

Reference: Last M. & Kandel M. (2001), Automated Detection of Outliers in Real-World Data, Proc. of the Second International Conference on Intelligent Technologies

# Pop-up menu

Selected options discussed above can be reached via the a pop-up menu which opens after a right mouse click in the table:

11079.93	1.26
73542.85	Reload
18569.45	Export Data
88223.70	Pivoting Customize
17968.00	Statistics
4628.00	Classify
20180.55	Table View
32530.40	3.75

The different options had been mostly discussed already above

- **Reload**: Loads the data anew from disk, useful e.g. when the files are re-written in the background.
- Copy to Clipboard: see "Clipboard export"
- **Export Data**: see "Export to file"
- **Pivoting**: see "Manually changing the pivot"
- **Customize**: see "Changing view options: fonts, number formatting and rounding, hiding empty cells, comparisons"
- Statistics: see "Adding statistics"
- Classify: see "Machine learning"
- View: see "Selection of the view type"

# Working with graphics

The exploitation tools allow showing the current numerical content of a view as a graphic. The underlying data are identical to those which would be found in a tabular view, however, in some cases, size restrictions will prevent to e.g. show all columns and rows. Most of the graphic types are based on the JFreeChart library (see <a href="http://www.jfree.org/jfreechart/">http://www.jfree.org/jfreechart/</a>).

While certain settings of the graphics can be changed such as fonts to some degree also colors, the graphics are not predominantly built in to provide camera-ready output for publications. Here, exporting the data e.g. via clipboard to a spreadsheet program is often the better option to control layout options.

# General handling of graphs

In the system, the selection of the view type including the different types of graphs is opened

View type

by pressing the view type **Table** bottom in the tool bar which shows the current active

view type (here "Table"). The following *graphic types* are currently supported, which are discussed in some detail in the following:

- Bar charts
- Line charts
- Area chart
- Waterfall chart
- Spider chart
- Pie chart
- Box and Whisker chart
- Histograms
- Scatter plots
- Deviation
- Network
- Markov chart

The *selection* of rows and columns shown in the graph can be set in three different ways, for all types of graphics:

• Using the **selection dialogs** (upper left corner of the table, or the buttons next to the graphic type selection drop down box: double click)



- Using **these filter buttons** in graphic mode: single clicks with the left mouse button will scroll down in the list, right mouse, single clicks will scroll up.
- Scrolling the table with the scroll bar to a specific position. The column/row in the upper left corner of the table will define the starting point for the graphic.

All types of graphics support *tooltips* to query the numerical values underlying the graphic. The tooltips appear when moving the mouse on a graphic element linked to the value as e.g. a bar.



Mouse clicks on the text of the legend items can be used to show **the numerical values** in the graph. A single left-mouse click on a legend item (here, e.g. on the text "Producer surplus, farm owned factors", adds the numbers to that series/block of value, while values for all other series/blocks become invisible. Pressing the "Control-Key" add/removes the values, the "Shift keys" affects all entries at once.



Perhaps an unexpected feature is the zooming in and out with the mouse. The graphs support, *saving* to the disk as a png file via a popup menu and *printing*. The popup menu also allows

changing certain properties for the current graph temporarily. Additionally, the export button can be used to save in different vector and bitmap formats to disk, see below.

Some settings which will pertain can be edited by opening the graphics option dialogue, press:

Ξ.	
2	-

Graphics settings		* [ Post		* : 100		×
Options for bar charts and histograms			Options for line and are	ea charts		
Maximal number of plots		4	Maximal number of plots	5		5 🌲
Maximal number of bar blocks		5 🌲	Maximal number of series			10 🌲
Maximal number of bars per blocks		10 🌲	Maximal number of observations			25 🌲
Foreground transparency in %		10 🌲	Foreground transparent	cy in %		0
☑ 3D effect	Plot vertical	Cylinder	3D effect	Plot vertical	Draw lines	Draw Shapes
Draw outline Draw shadow	Filled bars	Mean/median/q1/q3	Sort values	Draw cummulative	Common range fo	or plots
Options for spider charts			Options for pie charts			
Maximal number of axis		6 📤	Maximal number of plots	5		10 🔔
Maximal number of series			Maximal number of obse	ervations	25	
Maximal humber of series			Minimum percentage to	draw label		5 🌲
Foreground transparency in %		10 📮	Foreground transparent	cy in % 20 🚔	3D effect	
✓ Filled shapes			Labels inside of pies	;	Circular pie	
Options for Heat Maps, Histograms and Scatter F	Plots					
Show last column in Heat maps	📝 Show regression lin	e in scatter plots	📝 Show 50% predictio	on band in scatter plots	Show 75% predic	tion band in scatter plots
Show 90% prediction band in scatter plots	V Show 95% prediction	on band in scatter plots	👽 Show moving avera	ge in scatter plots	Show scatter plot	s above diagonal
Bins for xy-scatter (zero=default rule)		0 🌲	Dot size for scatter plots	s		1
Bins for histograms (zero=default rule)		0 🌲	Quantile around median			50 🌲
Options for all charts						
Font size relative to tables in %		60 🌲	Use shades of blue		Treat zeros as mi	issing values
Include zero in value axis range	V Show axis titles		Show domain grid line	es	V Show range grid	lines
		ok	Set color			

The chart type's specific settings are discussed in more detail below. The general options should be self-explanatory, it is best to try them out interactively.

## Setting colors manually

The "Set color"	button c	pens up	a dialogue
-----------------	----------	---------	------------

Graphics settings			Ireland
Options for bar charts and histograms Maximal number of plots Maximal number of bar blocks Maximal number of bars per blocks		Options for line and area ch 4 Maximal number of plots 5 Maximal number of series 10 Maximal number of observent	Arts
Foreground transparency in %	Stacked	Chose colors for graphics Color for series 1 Color for	
☐ Draw outline ☑ Filled bars Options for spider charts	Draw shadow	Color for series 5     Color for       Color for series 9     Color for series 9	
Maximal number of axis Maximal number of series		Color for series 13 Color for s Color for series 17 Color for s	
Foreground transparency in %	•	ok Color for series 21 Color for s Color for series 25 Color for s	Vorschau  Beispieltext Beispiel
Ontions for Heat Man		Color for series 29 Color for s	OK Abbrechen Zurücksetzen
Options for all charts Font size relative to tables in %	ast column		

With a button panel, which shows the currently active color ramp for the series. Pressing on of the buttons opens a color choser as shown above, which allows to change manually the colors used.

## "Walking" through the data

As the maximal numbers of elements shown is restricted (see above), typically not all columns and/or rows will be shown in a graph. The user basically has two possibilities to change the visible columns or rows. Firstly, columns and rows can be selected by the selection dialogues. Secondly, the user can click with the right or left mouse button on the buttons for table dimensions to move one row or column up or down.



### Exporting the graphic to file

The graphics can be saved to file in different formats by pressing the export button. The dialogue shown below will appear which allows the user to define the file and to select from different file formats. The "Options .." button opens an additional dialogue which is export type dependent, for PDF, to give an example, the paper size, orientation etc. can be changed. Generally, high quality are achieved if vector formats are used (PDF, SVG, EPS, EMF). These formats are however not supported by all applications, in opposite to e.g. bitmaps (BMP, which require a lot of disk space) and JPEG, which implies a loss of quality. It might therefore pay-off to try several formats for import into other applications. For MS Office users, the "Windows Enhanced Metafile" (EMF) format is interesting, as it allows changing later the graphics manually, e.g. by adding new text and changing colors.

The "Resolution factor" field allows to improve the quality of the saved file for non vector formats such as JPEG, by drawing the original graphic with more pixels which however drives up disk space.



# Exporting the graphic to clipboard

Alternatively, the graphic can be placed into the clipboard where it is stored as a bitmap or as

jpeg by double-clicking the "copy to clipboard" button.

# Pop-up menu in graphics

Selected options discussed above can be reached via the a pop-up menu which opens after a right mouse click in the table:

 Eigenschaften
 V
 Speichern unter
 Drucken
 Hineinzoomen
 Herauszoomen
 Autojustage
 Chose colors
 Customize graph
 Customize

As seen above, part of the pop-up menu, provided by the JFreeChart library, adjust to the language of the Operating System. The entries are briefly discussed below:

• Eigenschaften / Properties: Opens a dialogue where e.g. headlines can be entered, description changes and fonts selected for specific elements

Diagramm-Eig	enschaften		
Überschrift	Diagramm	Sonstiges	
Allgemein: Überschrift	anzeigen:		
Text: Schrift:		Arial, 7	Auswählen
Farbe:			Auswählen
		OK Abbrechen	

- Kopieren / Copy: see "Exporting the graphic to clipboard"
- Speichern untern / Save as: Open a dialogue which allows to store the graphics in png format
- Hineinzoomen / Zoom in: Opens a dialogue which allows to choose the axis for zooming in
- Herauszoomen / Zoom out: Opens a dialogue which allows to choose the axis for zooming in

- Autojustage / auto adjust: Reset the zoom to default.
- Choose colors: see "Setting colors manually"
- Customize graphs: see "General handling of graphs"
- **Customize**: general options to customize views, see "Changing view options: fonts, number formatting and rounding, hiding empty cells, comparisons"

## **Bar charts**

Bar charts assume that the columns – typically the table items – have different definitions and units, and consequently assign an own plot with a value axis to each column. The observations are taken from the table rows and define the domain, the horizontal axis. Each groups of bar columns present – typically the scenarios –receives its own colour. An example is given below.



GGIG Graphical Interface Generator - User Guide

Options for bar charts a	nd histograms			
Maximal number of plots				4 🌲
Maximal number of bar blocks				5 🌲
Maximal number of bars per blocks				10 🌲
Foreground transparency in %				10 🌲
☑ 3D effect		Plot vertical	Cylinder	
Draw outline	Draw shadow		Filled bars	Mean/median/q1/q3

The user has a number of options for the bar charts. By pressing the button in the toolbar, a dialog box including the section of "Options for bar charts" opens. The number of plots refers to the number of columns in the underlying tables, each column will receive an own plot with a matching value axis. The bar blocks refer to the rows, each bar block may comprise several bars taken from the column groups (typically scenarios). As seen above, it is also possible to generate stacked bars from the column groups, or to generate cylinders instead of cubes.

## Waterfall charts

A waterfall chart is a special version of a bar chart which shows from left to right (in the underlying table from the top row to the bottom row) the transgression from a starting value to a final value. Negative elements are shown in red, positives ones in green. A possible application is shown below which depicts how policy instruments and costs of young animal define the producer incentive prices from the world market price. If the last value is not approximately equal to the sum of the other, an additional observation with the sum is added automatically.



## Line, point and area charts

Line and point charts assume that the columns of the table present some ordered sets e.g. years or iterations. There is currently a default of 25 such observations which can be increased by the user. The different series to plot are taken from the table rows. If different column groups are present, those receive their own plot with an own value axis.



The options for line and area charts are similar to the ones for bar charts. The number of plots refers to the column groups, the number of series to the rows of the table.

Area charts are equivalent to stacked bars, i.e. the observations are added to each other. The number of observations is linked to the columns. Negative values are shown below the zero axes. An example was shown above.

### **Pie charts**

Pie charts are useful to show shares on total as e.g. trade flows. The shares are calculated from the columns, whereas each column group – typically scenarios – receives its own pie. Only one row is allowed.



The user has the following options to modify the presentation of pie charts:

Options for pie charts						
Maximal number of plots	4 🗘					
Maximal number of observations	25 🛟					
Minimum percentage to draw label	100 😂					
Foreground transparency in %	20 🛟					
☑ 3D effect	Labels inside of pies					
🗹 Circular pie						

The maximum number of plots refers to the number of elements in the dimensions of the column group., The example above shows two plots. The number of observations defines the

numbers of pies – if more columns are available, the cake will eventually give a wrong impression if not all values are used to define the sum and the shares.

The "minimum percentage to draw label" defines a lower cut-off limit, if a cake's size is below the threshold, no label will be drawn. As shown in the example above, setting the threshold to 100% will erase the labels (see Pie chart maps for an example). It is also possible to place the labels in the pies, and not outside of the cake as shown in the example above.

# **Spider plots**

Spider charts are useful to compare several dimensions simultaneously across a range of alternatives. It is assumed that the columns show the items of which each receives its own axis, whereas the column groups are the alternatives to compare. The axis are not ticked with numerical values, instead they are always scaled to cover the minimum and maximum found in any alternative.



GGIG Graphical Interface Generator - User Guide



The options for spider charts which are found under the *button* in the toolbar are rather limited. The user can determine how many axes – taken from the columns – are included in the diagram and the maximum number of series, which typically consist of scenarios.

# **Box and Whisker charts**

In descriptive statistics, a box plot or boxplot (also known as a box-and-whisker diagram or plot) is a convenient way of graphically depicting groups of numerical data through their fivenumber summaries (the smallest observation (sample minimum), lower quartile (Q1), median (Q2), upper quartile (Q3), and largest observation (sample maximum). A boxplot may also indicate which observations, if any, might be considered outliers (see also <u>http://en.wikipedia.org/wiki/Box\_plot</u>).

Boxplots can be useful to display differences between populations without making any assumptions of the underlying statistical distribution: they are non-parametric. The spacings between the different parts of the box help indicate the degree of dispersion (spread) and skewness in the data, and identify outliers. Boxplots can be drawn either horizontally or vertically (text so far from Wikipedia).

The box and whisker chart uses the rows as the observations, and generates an own graph per column. The box shows +/-25% of the observations around the median which is shown as a grey line, whereas the arithmetic mean is shown as a grey circle. The whiskers show the median +/- three times the inner quartile range. Mild outliers are drawn as dots and strong outliers are indicated by arrows. So far there are no specific options for that type of diagram.

💰 Exploit scenario results [	[0							
Table	Production activity	Item		Years		View type	0	•
Supply details, mapping view	Cereals	✓ Income [Euro/ha or hea	sd]	✓ 2013 ✓	∎ ➡ ◄ 🔛 롣	Box and Whisker chart	Ţ	
2300.00					0			
2100.00								
2000.00								
1900.00								
1800.00								
1700.00								
1500.00								
1400.00			-					
1300.00								
1200.00								
1100.00								
800.00 -								
800.00								
700.00								
600.00 -			1		•			
400.00	•							
300.00			]					
200.00								
100.00								
0.00								
-200.00								
-300.00								
L								
		г	Scenarios					
			BIOF_D2E2 BIOF_D10E1	10				

# Histograms

As for whisker charts and statistics shown in tables, the observations are taken from rows, and different columns are charted individually. Please note that it is also possible to generate a separate Histogram window, but then, the observations refer to all columns simultaneously.



Some tips:

- If the data set comprises zeros which should be interpreted as **missing values**, check the box "Treat zeros as missing values". Otherwise, the value axis will show a zero as the lower bound even if "Include zero in value axis range" is not ticked.
- The **number of bars** (the so-called bins) can be set with a spinner in the second lower panel in the graphic dialogue. The zero as the default value determines the number of bins automatically as the minimum of 100 or the number of observations divided by 5.
- It might be worth to increase with t**ransparency** of the bars to better capture overlapping parts of the distribution. It might also be worth to use unfilled bars.
- In order to draw a **continuous distribution** instead of bars, set "draw outline" and "Filled bars" to off, and use "Draw line" which generates a graph as seen below
- With "Show mean/median/q1/q3", switch on the marker lines (mean: normal line; median: dotted; q1/q3: dashed) as seen in the graph below.



As the continuous distributions are drawn with a spline renderer, they can be quite nicely smoothed if the number of bins is decreased (in the example below from 1000 to 100 to 10 bins):



### **Scatter plots**

Scatter plots are an extension of Histograms, showing the one-dimensional distribution of items on the diagonal panels as histograms, their combined distribution on the off-diagonal as a scatter plot below the diagonal. The plots above the diagonal generate equally sized cells in x and y, and indicate by the intensity of the color how many observations are found in each cell.

The scatter plots comprise (on demand) a linear regression line through the data with a  $\pm$ -5% prediction band, and a moving average line using 10 observations. All elements are switched on in the example shown below.



The histograms can be customized as discussed in the section above; the way the observation is organized is identical to histograms as well. For the scatter plots, different options can be changed:

Options for Heat Maps, Histograms and Scatter Plots							
Show last column in Heat maps	Show regression line in scatter plots	Show 50% prediction band in scatter plots	Show 75% prediction band in scatter plots				
$\fbox$ Show 90% prediction band in scatter plots	Show 95% prediction band in scatter plots	Show moving average in scatter plots	Show scatter plots above diagonal				
Bins for xy-scatter (zero=default rule)	0 💌	Dot size for scatter plots	1				
Bins for histograms (zero=default rule)	0 🛋	Quantile around median	50 🔦				

1. Changing the number of scatter plot bins affects the number of xy-cells above the diagonal; a value of 5 generates 5x5 = 25 cells and gives a graph as follows:



2. Changing the dot size can ease the visualization of the distribution:



3. Finally (not shown here graphically), the regression lines, moving averages and different prediction bands can be switched on and off.

In the context menu of the graphics,



The weight of the current sub-plot can be increased and decreased (in several steps), to yield a layout as in the example below:


## **Deviation renderer**

A deviation renderer shows the median of a series along with a symmetric quantile around it. That makes only sense if one has an ordered series, e.g. points in time.



In order to produce such as graph:

- Put the observations (e.g. years) in the columns
- Put the draws in the rows

• Put the scenarios in the column groups

The bandwidth of the graph can be changed with the "Quantile around mean" setting in the graphic dialogue, below the bandwidth is 22%: 11% above and 11% below the median.

Options for Statistics	
Bins (zero=default rule)	0 🛬
Quantile around mean	11 🔹

The deviation renderer can be expanded to resemble a **contour plot** which shows the density inside the band by lines which combine point for each observation of equal cumulative probability, if the draw line shapes is switched on:



## How to draw a line chart with mean / min / max etc. over a time series

Alternatively to a deviation renderer, one can also show more statistics as time series. In order to do so:

- Put the observations (i.e. the stochastic draws) into the rows of a table
- Put the time dimension in the columns
- Use the pop-up menu in the table under "statistics" to show the statistics you are interested in the table

Region	w, uraws (0)			Activity			Product	
Lake Nava	ia			Reservoir			Levi Im	above se
<b>1</b>	2	3	; 4		5	6	7	8
d1	1887.22	1886.60	1885.96	1885.36	1884.80	1884.30	1883.85	
d2	Reload	ł	1886.98	1887.24	1886.43	1885.43	1885.39	
d3	Copy	to Clipboard	1887.70	1887.67	1887.72	1887.41	1886.75	
d4	Export	Data	1887.45	1887.52	1887.45	1886.91	1885.50	
d5	Diveti		1889.03	1888.23	1888.23	1887.32	1886.39	
d6	Fivoti	· · · · ·	1888.30	1888.54	1887.87	1886.53	1886.07	
d7	Custor	mize l'able	1885.80	1885.56	1885.88	1886.13	1886.68	
d8	Statist	ics	1886.91	1886.93	1886.85	1886.93	1886.35	
d9	Classif	iv.	1886.92	1886.98	1886.58	1885.71	1884.75	
d10	Classii	у	1886.25	1885.01	1884.90	1885.45	1884.95	
d11	View		1888.40	1887.84	1888.18	1888.81	1888.33	
d12			1886.97	1886.29	1885.86	1886.50	1886.50	
d13	Table	View	1887.26	1888.27	1887.85	1886.50	1885.70	
d14	1887.49	1886.48	1885.78	1885.77	1885.47	1886.44	1886.29	
d15	1887.49	1886.49	1886.63	1888.38	1888.21	1888.10	1886.73	
d16	1887.44	1886.70	1886.04	1886.06	1886.38	1886.28	1886.86	
d17	1887.66	1887.78	1886.97	1886.42	1887.85	1886.81	1887.87	
d18	1886.99	1886.80	1887.06	1887.40	1887.45	1886.71	1886.03	
d19	1888.64	1887.80	1886.66	1885.83	1885.22	1885.55	1884.59	

1. Add the stastistic which you want to show:

٠

🛓 LANA - HERE	AMO [T:\britz\lan	a]						
View Handling	Windows							
🙆 Reservoir o	verview, draws [0]							
Reg	ion			Activity			Product	
Lak	e Navaja			Reservoir			Levi [m	above sea levi]
	1	2	3	4	5	6	7	8
		-	Č.		, and the second	°		Ĭ
÷	1887.2	2 1886.60	1885.96	1885 36	1884.80	1884 30	1883.85	1883.4
<u>d1</u>	1887.5	3 1887 68	1886.98	1887 24	1886.43	1885.43	1885 39	1884.6
d2	1887.3	2 1888.44	1887.70	1887.67	1887.72	1887.41	1886.75	1889.3
d4	S	et statistics	-	-		×	1885.50	1884.6
d5						2	1886.39	1887.3
d6			C On	v show outliers		3	1886.07	1886.4
d7				,		3	1886.68	1886.7
d8			Set factor for r	nean/sigma	2 🌲	3	1886.35	1886.1
d9			Set maximum perce	ntage of outliers	2 📥	1	1884.75	1885.4
d10			See maximum perce	rage of oddiers	2 💌	5	1884.95	1884.8
d11	Se	t outlier detection m	ethod Standard de	viation around me	an	- 1	1888.33	1887.6
d12	Sele	ect statistics				0	1886.50	1885.9
d13	Not	)S				0	1885.70	1885.4
d14	Mea	an tiao				4	1886.29	1887.1
d15	Std	Dev				0	1886.73	1886.8
d16	<b>q1</b>					8	1886.86	1887.0
d17	q3					1	1887.87	1887.6
d18	min					1	1886.03	1886.2
d19	ma	c Outlier				5	1884.59	1884.3
d20	max	Outlier				3	1887.70	1887.2
d21				ancel Unda	te	8	1884.60	1885.6
d22				uncer Opua	œ	0	1886.64	1885.5
d23	40000	4000 50	4000 50	4007.07	4000.40	4850.04	1887.57	1887.1
d24	1886.9	1886.53	1888.53	1887.87	1888.46	1888.04	1888.14	1889.3
d25	1008.4	0 1008.20	1008.33	1068.23	4994 72	1000.01	1000.93	1000.9
d26	1887.6	6 1887.96	1887 84	1887 48	1886 78	1885 70	1885 25	1885.6

2. Switch to line chart views

Attention: the statistics do not work correctly if several dimensions are merged in the rows.



#### GGIG Graphical Interface Generator - User Guide

3. And use the graphic dialogue to select as many series as statistics selected (in the above example 4)

## **Network graphs**

Drawing on the Jung Graphics library (<u>http://jung.sourceforge.net/</u>), network graphs visualize vertexes and edges of directed graphs, e.g. to depict the flows in an economy. The following screen shot shows an example from the GTAPinGAMS project depicting income generation and use:



The graphs can be customized in different aspects:

Graph layout Attraction 0,75 Repulsion 0,75 Zoom + • • Ø zoom at mouse (wheel only) Font Ø bold text Font size relative 60 Mouse Mode TRANSFORMING Label Pocifico	Edge shape Ine wedge Quad curve cubic curve orthogonal fill edge shapes show edge values Show arrows V directed V centered	Nodes v color by change stroke highlight on selection v show values shape by degree v size by value stretch by degree ratio filter when below min degree Min degree
	Store layout Remov	re layout

Most of the options are best tried out interactively on a graph. "Store layout" fixes the positions of the vertexes during subsequent updates of the graph, "Remove layout" let the underlying graphing engine determine positions again automatically.

The underlying item definitions in the view must define the edges.

## **Markov charts**

A still explorative type of graphics visualizes flows between entities which are placed in a two dimensional co-ordinate system. It is currently not yet used in CAPRI itself, but applied to

show flows between farm groups classified by economic size and specialization. As with the flow maps below, the major code based for the graphics is based on work of Doantam Phan<sup>3</sup>.

The positions on the x and y co-ordinate are deducted from the codes, taken from a specific section of the underlying XML-definitions which refers to a matching of sub-strings of the codes and x respectively y positions. The size of the dots is taken from the diagonal elements.



# Working with maps

## Flow maps

Flow maps visualize flows between regions - the maps are constructed by taken the elements in the rows as the origins of the flows, and the elements in the columns as the destinations. Flows from the same origin are drawn in the same color, the width of the flows relates to their size. Counterfactuals are taken from the column groups and receive a specific "dash". The picture below shows a screen shot of a flow map for two scenarios.

<sup>&</sup>lt;sup>3</sup> Flow Map Layout, Doantam Phan, Ling Xiao1, Ron Yeh1, Pat Hanrahan, and Terry Winograd, Stanford University, see <u>http://graphics.stanford.edu/papers/flow map layout/flow map layout.pdf</u>. I would like to thank Doantam Phan for letting the CAPRI team use and modify his source code.



When pressing the map option button **!**, the following dialogue is opened:



The main options of interest for flow maps are the scaling model and the display width. The following *scaling models* are available:

- *Linear*: the width is determined by relating the flow quantity to the sum of all flows for the same scenario.
- *Log*: the width is determined by multiplying the log of the relation between the flow quantity and the minimal flow with the log of the relation of the maximal and minimal flows for the same scenario.
- *Polynomial*: the relation between the current flow and the maximal flow is raised to a power determined by taking the log of the relation between the maximal and minimal

display width divided by the log of the regional between the maximal and minimal flow.

The user can prevent that small flows are drawn by setting a minimal width relative to the size of the window; equally, the maximal possible size of a flow relative to the size of the window can be determined.

In order to show only a selection of the flows, the selection 😨 🖻 buttons can be used. The lower left one relates to the rows of the underlying tables, and thus allows excluding origins from the maps. The lower right one opens a dialogue to exclude destinations, whereas the upper right one allows exclusion of scenarios.

Most options described below for thematic maps such as zooming and dragging are also available for flow maps. However, classifications and color models cannot be supported.

# Pie chart maps

Another rarely used application of maps is the possibility to place pie charts above the geometry. The regions must – as always with maps – be placed in the rows of the underlying tables, and the cakes are calculated from the data in the columns. It is possible to produce maps for different scenarios when those are placed in the column groups as shown below. The size of the charts depends mainly on the bounds of the underlying polygon, so that smaller countries have smaller pies. The settings for pie chart diagrams (see Pie charts) can be applied to that view.



## Colored thematic maps

For CAPRI, the GUI currently provides geometries for NUTS 2 regions, Member States, the regions with behavioral functions in the market model, trade blocks in the market model and finally, the Homogenous Soil Mapping Units (1x1 km resolution) underlying the spatial down-scaling component. The geometries are always linked to the rows of the underlying table.

The most obvious way to visualize results is the use of thematic maps. This holds true for NUTS2 results, but even more so for the results at the HSMU level. When starting the GUI, the mapping view uses some pre-sets which can be interactively changed as described below. The following screen-shot shows the result of first loading the base year results from the spatial dis-aggregation for Denmark and then switching from the tabular to the mapping view.

As with other views, the content of the map can be changed by working with the drop-down boxes, or by (de)selecting columns and rows. There are specific possibilities to change class limits, colors and further features for maps which are discussed in the following.



## Changing the classification and the legend

In order to change the layout of the map, click the mouse in the area of the legend or double-



It offers different options to change the way the map is drawn on screen and information

supporting the classification.

## Changing the coordinate set



The map viewers assume that the regions are the in rows and searches the short keys for each row in the ids of the coordinate set. Regions for which no matching id is found cannot be drawn. If, to give an example, a coordinate set of EU Member States is used and the table comprises data at NUTS2 level, only those NUTS2 regions will be drawn which are identical to an EU Member state. In that case, one needs to switch to the "NUTSII.zip" coordinate. In most cases, the XML file which defines the view will already comprise the information about the appropriate coordinates to use.

The following message will appear if for none of the items in the rows a matching id is found. That is most often the case if the regions are not shown in the rows.



The coordinate sets are typically stored in the GUI directory. The map viewer uses a proprietary format which aims at minimizing disk space. A utility (see section on utilities) allows generating the proprietary format based on a so-called shapefile, a format widely used in GIS work.

## Adding a histogram window to a map

In the map option dialogue, tick "Show histogram" and a separate window with a Histogram will be shown. It will use the current classification and color model to visualize the distribution of the values, reports some basic statistics and shows a box and whisker diagram.



## Shrinking polygons according to UAA share

The optical impression received of a map where colors are used to distinguish between values, depends to large extent on the area covered by a certain color. If the majority of the pixels is drawn in red, that will send a warning message to the user. In the case of the HSMUs and information relating to agriculture that message can be strongly biased as almost all HSMU comprise some other land cover then agriculture, and some of the HSMU comprise only very little agriculture, but e.g. forest, shrub lands, water bodies or artificial land cover. The HSMU geometry therefore comprises the information about the share of UAA assigned in the base year to each HSMU. That information can be used to shrink the area of the polygons when drawn on screen accordingly. That is done by drawing all points of the polygons towards the centroid of the polygon and then multiplying the distance between the point and the centroid with the square root of the share of the UAA. In the original HSMU geometry, such polygons had been broken down to simpler ones where the connection between a point and a centroid would cut through a segment of the polygon. In such cases, shrinking could let the new polygon hide other ones.

The graphs below show the very same map (same input data, classification and coloring) for the High Nature Value indicator for a part of Belgium. The right hand side map draws the HSMUs into their full size, the one on the left hand side one uses shrinking. The message perceived is probably very different. In the unshrinked right map, one may conclude that there is a lot of highly intensive agriculture (low HNV indicator drawn in red) in the lower diagonal triangle and some important areas of high nature farmland in the protruding area. This optical impression differs strongly from the polygons drawn with corrected shares for agricultural cover. It turns out that in the lower diagonal triangle, the density of agriculture is often low, and especially low in the intensively managed HSMUs. Equally, it turns out, that the area covered by High Natural Farmland in the protruding part is relatively small.



## Area weighted classification

The classification can be generally applied treating each "region" (e.g. a NUTS II or a HSMU) as an observation with equal weight or using the areas of the underlying polygons as weights. Those weights are multiplied with the share of UAA if shrinking is used as explained above.

## Excluding zeros from classification and removing small and large values

In GAMS, zeros and missing values cannot be distinguished. For certain results, zero results are therefore coded as very small numbers to allow for that distinction. Zero observation can be excluded from classification and the polygons with zero observations will not be filled. Equally, a number of regions with small and large values can be excluded from classification which might help in finding an appropriate coloring scheme and legend.

Number of regions with small/large values excluded from classification 0 🖨 0 🚔

## **Classification method**

The "classification method" selects the ways the class limits are defined, i.e. the lower and upper bounds for the numerical values drawn in the same color. Generally, the same quantitative information might generate quite different graphical impressions depending on how classes are defined and colored.

For all types of automatic classification methods a clean-up procedure is used which removes classes with identical limits. It is generally recommended to use a number of classes which can be easily identified by the user, and to consult the frequency or cumulative distribution graphs present in the map option dialogue to check to what extent the class limits chosen represent the data well. It is generally recommendable to try out different classification.

The following classification methods are currently supported:

#### Natural breaks

Natural breaks classification is a method to cluster the data into classes so that differences between the means of the classes become high while the standard deviation inside the classes becomes low (FISHER, W. D. (1958). "On Grouping for Maximal Homogeneity", Journal of the American Statistical Association 53, 789-798. Code based on: HARTIGAN, J. A. (1975). Clustering Algorithms, John Wiley & Sons, Inc, New York, pages 130-142.). The algorithm does not only find the approximate best solution, but often gives rather appealing class limit definitions.

It works rather well if no extreme outliers are present in the distribution. In the latter case, classes solely comprising the outliers will be generated, and the vast majority of the values will be put in one or two classes. Here, it might help to exclude some very large and small values from calassification.

The clustering algorithm is rather expensive to calculate, so that in cases in which the population exceeds 500 observations a somewhat simplified version is implemented in the CAPRI GUI. From the original observations, a "condensed" population is generated whose members represented means of consecutive observations of the original one. The members are set so that the number of observations from which the mean is calculated is not bigger then 1/500 of the original population size and that the spread of those observations is smaller than the minimum of 1/500 of the spread of the total population and 10% of the standard deviation. The actual calculations are then done taking the size of the resulting classes into account.

## Quantile

The observations of the regions are split in a way so that approximately the same number of observations fall into each class . Quantiles are cheap to calculate and are therefore the default setting, and often appealing as colors occupy similar areas in the overall map as long as the polygons have approximately the same size.

If unique values are found at the end of a quantile, the algorithm will either exclude all observation with that unique value from the class or include all of them. The decision will be based on the fact if with or without inclusion the size of the class comes closer to the desired size. If the user has e.g. chosen five classes, the desired class size should cover 20% of the observations or area weights.

Quantile based classification work not well if the distribution is far from uniform in which case the differences between the classes might be rather different.

#### Equal interval

The differences between the current minimum and maximum value is divided into classes of equal spread. This may lead to rather curious class limits when outliers are present. In those cases, it may be appropriate to exclude some regions from the classification. See below for details how to exclude regions from the classification.

#### Mean standard dev

The class limits are defined according to the mean and the portions of the standard deviation of the data. It works best with normally distributed data, but may result in very small classes if the distribution is skewed, e.g. long tailed. The algorithm will always introduce at least four classes, then six, eight, ten and twelve. More than twelve classes are neglected.

The algorithm takes into account the spread of the data, and sets the class limits accordingly. If all observations fall into  $\pm -25\%$  of a standard deviation, class limits are introduced at 25% and 10% for four classes. If the number of classes is higher, new limits are introduced at 5%, 2.5%, 1% and 0.5%. In case of  $\pm -50\%$ , the smallest class is dropped and  $\pm -50\%$  added, and so forth up to  $\pm -3$  standard deviations.

## Nested mean

The nested mean classification will only work with 2, 4 or 8 classes. The classes will be defined such that one break is found at the mean of the sample. The resulting two halves of population are then again divided by their mean to get four classes, and the resulting quarters divided by their means to define eight classes. This works well with rather skewed distributions.

#### Manual classification

Finally, the user may set the class limits by hand. In order to do so, double click the mouse on the appropriate row in the table with the classification results in the column "class limit". The value can now be changed with the keyboard. When this is done, click into another cell. The labels will be adjusted accordingly. Afterwards, when all class limits are defined, the user may also overwrite the label (e.g. using words as "low" or "high").

Please keep in mind that currently the values will be lost if you load other data or change the classification, number of classes etc..

#	label	class limit	% of obs	color
1	0.00 < 0.00	0	30.488	
2	< 28.41	28.407	17.398	
3	< 51.86	60	17.398	
4	< 112.82	112.821	17.398	
5	< 586.84	586.835	17.317	

#### Integration distribution information in the map window

The GUI allows the user to enter distribution information in the map in different ways. The first possibility is to print a simple frequency diagram above the legend.



That gives a rather intuitive feel on how well the class limits represent the data distribution. In our example below, it is obvious, that the majority of the values lie in the first class.



Less suitable for final out, but useful while playing around with classification methods and class definition are the distribution dots which can be added. They carry additional information on the location values in different classes.



Finally, switching to linear or logarithmic may be a way to help reading the map.



## **Color table**

The color table or color map defines the colors used for the classes. When choosing the color model, keep in mind that colors carry a meaning; red e.g. is generally interpreted as dangerous. Equally, it is important to think about the final medium with which the map will

be published. Exporting colored maps to a black-white device will render it almost impossible to read the map. It is best to try different color tables and different classification methods on your data. The following color models are currently available, named according to the data order from minimal to maximal value:

- Green Yellow Red (standard): Normally, the middle class is drawn in yellow, smaller values in shades between yellow and green, and larger ones from green to red. This should be applied e.g. to environmental indicators where the damage increases with the value of the indicator.
- *Red Yellow Green:* as above, only that high values are shown in green. Should be used e.g. for income indicators or environmental benefits.
- *Red Gray Green / Green Gray Red:* more available for historic reasons as they mimic the color tables of the original JAVA applet.
- *Blue Gray Green / Green Gray Blue:* introduced on demand of DG-AGRI. A good choice if the "good"/"bad" interpretation of the distribution is to be avoided.
- *Shades of grey:* sometimes needed for publications when color printing is not available in the final hardcopy. Beware to use a limited number of classes.
- *Shades of blue*: useful where the notion of "bad" or "good" inheritably comprised in greenish and reddish colors is to be avoided.

If percentage or absolute difference or some other quantitative range is sued which has a natural middle point (for difference that the zero case, i.e. no difference), one should use

```
Set value for middle color 950.81 👻
```

to chose the class closest to the natural "middle point" for the middle color, see below.

#### Defining an self-created color model

Once a color model is chosen, the user can re-define the start, middle and end color using the three buttons on the color table selection row, as shown below, given a lot of freedom to generate color ramps.



#### Changing the value for the middle color

Normally, the medium color (yellow or gray) is assigned to the middle class. Sometimes, the user may wish to change the class where the color switches. First, the "Set value for color change" must be ticked. Next, in the now enabled drop-down box, choose the class limit for which the middle color should be used. The effect is shown below. Before, values in the class below "392.70" – the middle class – were drawn in yellow. When the user now selects another class limit, the colors assigned to the classes change. Here one of the shades of green is dropped and shades of red are added.



#### Manual set colors

Finally, the user can choose its own colors by double clicking in a color field in the legend table. That should only be done after the final definition of the class limits is set as otherwise, the manually set color will be lost.



## Changing the way the legend is drawn

The map viewer always puts the legend below the map. Currently it offers three options how legends are drawn:

1. Separate, equally sized rectangles which show the upper class limit with the exemption of the lowest class, which shows the lower limit.

Y

	•	-					
> 00.0	< 49.28	< 59.92	< 74.08	< 80.19	< 86.45	< 113.44	< 450.86

2. A continuous linear scaling bar. That gives an optical idea about the distribution of the class limits. Overlapping of the number is avoided by skipping class limits close to

Legend: Separate rectangles



In all the cases, the tool dialogue can be used to set number of digits shown, e.g. reducing the number of digits to zero leads to a linear bar as shown below:



The reader is reminded that the label can be changed manually as shown below.



## Changing the title of the map

When using output to clipboard or disk, the user may often prefer to choose his own title or no title at all on top of the map. This will be helpful when producing a caption for the map in another application. In order to refrain from drawing a title on top of the map, click into the legend part of the map, and in the dialog at the bottom, choose "none" in the row labeled "Title on top of map". Alternatively, the user can simply write something in the box.

Standard map title	
No map title	

## Zooming in and out and navigating in the map

In order to zoom in part of the map, press the 🔍 button. The mouse pointer will change to a magnifying glass with a cross in it. You can then mark an area on the map by pressing the mouse button, dragging and then releasing the mouse. After the mouse is released, solely the selected zone of the map will be drawn, without changing the class limits or any other setting. Clicking with the mouse while being in zoom in mode will increase the map resolution stepwise by 25% and center the map at the current mouse position.



By clicking with the zoom out pointer  $\square$  on a point of the map, the point becomes the new center point of the map and the map resolution is reduced stepwise by 25%. Equally, you may drag the map while keeping the current resolution by choosing the drag pointer  $\oplus$ . Finally, in order to return to the original full-sized map, use the "full extent" button  $\oplus$ . The reader should note that the "full extent" button shows a rectangle around the arrows.

## Getting data for specific polygons

The info pointer 1 will open an additional window, as shown below, which displays information on the current polygon – the circle above the "i" being the focus point. The title bar of the new window shows the code and, if available, the long text of the polygon currently pointed to with the info pointer. The content of the info window is continuously updated when the mouse is moved over the map, and all polygons belonging to the same region as the one pointed on with the mouse is highlighted.



If the user opts to use one of the comparison options to be shown (percentages, differences, normalization) by clicking on the "customize" button,

Arial	• 11 • plain •					
Frection digits and dec	imal separator 2					
Separator between me	Separator between merged data dmensions					
Column width 750	Column width 750 1 Row width 750 1					
E Hide empty rove	s 📉 Hide empty columns					
Cut off levit to determine	interior aller					
	authol care a (a)					
Vise default pivoting for tables Show	e histogram					
Use default pivoting for tables Show	Hostogram Use classification colors for tables Using texts only					
Use default pivoting for tables Show Show only selected items Comparison sulput	Histogram Use classification colors for tables Use classification colors for tables Only texts only Only values					
Use default pivoting for tables Show Show only selected items Comparison output Data dimension used for comparisons	Herdy dass Herdsselfcation colors for tables Long texts only Chrly values Dry values					
Use default pivoting for tables Show Show only selected items Comparison output Data dimension used for comparisons Bernent used for comparisons	Histogram Cong texts only					
Live default pivoling for tables  Show only policited items Show only policited items Data dimension used for comparison Benerit and for comparison or, define calors define table	whethy the  we describe a colors for tables we describe a colors for tables we describe any we describe a colors of the colors color advance colo					
Use default pivolog for tables Show Show only selected items mpartice autour to diversion used for conparisons ment used for cosparisons ok define calors define ate	Conjecture and a second s					

an additional column is automatically added to the info window showing the comparison value used. That is especially helpful when the map shows only differences.

The content shown in the info window is not fixed, rather, the user can decide which data dimensions to use for the columns and rows by using the "map option dialogue" by clicking on the legend of the map. If the user e.g. switches to "items" instead of "activity", the "info" window will look like shown below. An alternative is to use a second tabular view in addition to the map.



## Highlighting specific regions in the map

Sometimes it may be interesting to see the spatial distribution of specific data or data constellations. All views open the possibility to (de)select columns and rows, allowing e.g. to use the NUTS code in front of the numerical HSMU code to select only the HSMU belonging to specific administrative regions. That possibility is explained in short. First, double-click the row selection button II ("Open selection dialog for table rows") which will open the following dialogue.

Selection dialog for Table rows		X
Enter search pattern in field and us	e bul	ttons, or use mouse to define selections
DK001_2_3*		
Clear selection, add pattern to label	s	Clear selection, add pattern to keys
Add pattern to Clear list of s	selec	ted items, and define the selected items according to
Remove pattern from labels		Remove pattern from keys
DK001_2_3_4H22676		H22508
DK001_2_3_4H22677		H22509
DK001_2_3_4H22678		H22510
DK001_2_3_4H22679		H22511
DK001_2_3_4H22680		H22512
DK001_2_3_4H22681		H22513
DK001_2_3_4H22682		H22514
DK001_2_3_4H22683		H22515
DK001_2_3_4H22684		H22516
DK001_2_3_4H22685		H22517
DK001_2_3_4H22686		H22518
DK001_2_3_4H22687	-	H22602
ОК		Cancel

Now, we may e.g. select only the HSMU belonging to the FSS region DK000\_1\_2\_3 by typing ""DK001\_2\_3" in the left input box, and then choosing "Clear selection, add pattern to labels". Afterwards, the map will look as shown below.



The tabular view opens up the possibility of using numeric filters, an option discussed in the following. Take for example the task to select all regions where the Nitrogen Fertilizer Consumption is between 100 and 150 kg/ha. First, switch from map to tabular view. In the

table click with the right mouse button in the column header of that column holding the values to which the filter should be applied, as shown below. We will need to apply the filter stepwise, first e.g. selecting all values greater than 100 and then removing those which are above 150.

lection	View Handling Windows	
Build database	🚔 Exploitation of spatial results [Data View 1]	
Generate baseline	Activity Items	Table
Edit simulation	UAA 💌 No 5 (a), Mineral Fertilizer Consumption, Nitrogen kg N/ha	Agri-Env indicators, driving forces
Run simulation		BASE A
Ex-laib-labe-base-service	×	DHOL
Ex Define numerical selection filter	for table rows	
Ex Comparison operat	or Comparison value	
Ex > 💌	100	
De Clear selection and select acco	rding to filter	
Add result of filter to existin	g seleStion	
Remove result of filter from exi	ting selection	
OK Cance		
	H22515 H22516	

After clicking on "clear selection and select according to filter", and then on "ok", the table will only show such regions where the value in the column "BASE" is above 100, as shown below. Next, we must exclude the regions above 150 kg/ha. To do so, set the filter to ">" "150"m and then press "remove result of filter from existing selection".

Ger	nerate baseline Activity Iter	ns	
i Ed	∯ Filter dialog	×	n kg N/ha
1 Ru	Define numerical selection filter for table rows		
1 Ex	Comparison operator	Comparison value	
Ex	> 💌	150	147.70
Ev	Clear selection and select according to filter		169.36
, LA		1	176.21
) Ex	Add result of filter to existing selection		164.02
1 De	Remove result of filter from existing selection		99.38
			95.50
	OK Canta		130.30
			99.90
	H2873		160.19
	H2874		134.93
	H2875		29.19
	H2876		77.86
	H2877		126.38
tatio	n mode selection H2878		113.30

Now, drawing a map with just those regions is not so interesting. However, with the tool dialogue, we can highlight the selected value instead of hiding all others. The selected rows are now shown in red in the tabular view.

		É Exploitation of spatial results [Data View 2]	
		Activity Items	Tab
Customize view	٦	UAA 💌 No 5 (a), Mineral Fertilizer Consumption, Nitrogen kg N/ha	▼ Agr
	Ш	BASE	
Arial v 11 v plain v		H2865	147.70
		H2866	169.36
Fraction digits and decimal separator 🛛 🗸 🗸		H2867	176.21
		H2868	164.02
Separator between merged data dimensions 👻		H2869	99.38
		H2870	95.50
Column width 750 A Row width 750 A		H2871	130.30
		H2872	99.90
		H2873	160.19
Hide empty rows		H2874	134.93
		H2875	29.19
Cut off limit to determine empty cells		H2876	77.86
		H2877	126.38
I Use default pivoting for tables		H2878	113.30
		H2879	61.16
Channell Harris Michilla Harris and a large starting and		H2880	60.65
Show all items, highlight selected		H2881	241.39
Comparison output Show only selected items		H2882	241.38
Show all items, highlight selected		H2883	124.49
Data dimension used for companisons Region -		H2884	108.58
Element used for comparisons France 👻		H2885	105.87
		H2886	73.94
ok define colors define statistics store settings load settings		H2887	136.51
		H2888	83.92
		H2889	108.16

When we now draw the outlines of the selected polygons only (see map option dialogue), the map will draw the outline of the selected regions in cyan and thus highlight them. The row selection will be maintained when the pivot or the table is changed, as long as one of the selected items can be found in the rows of the new table. The example map shown below is certainly not so interesting, as changed class limits could have done basically the same job. However, we could switch e.g. to grass land shares to see if fertilizer input is more often found on arable or on grass land.



## Updating the map

Generally, the map is updated automatically when the user changes an option with an impact on its layout, as long as the number of visible polygons is below 20.000. If that amount is exceeded, the classification dialogue is updated immediately, but not the underlying map. In order to apply the changes, the "apply" button must be clicked on. The user is informed that the "ok" button will also update the map, so that an "apply" immediately before an "ok" is not necessary.

## Adding region label to the map

In the map option dialogue, tick the box "Show regions labels in map"

Region labels	Label options
N .	

to add labels to the largest polygon for each region as shown below.



By clicking on the Label options ... button, the "Region label steering" dialogue box opens which allows changing some settings. For maps with just a few regions (or when zooming), it might be worthwhile trying to play around with the action to improve labeling.

🕌 Region labels steering		×
Font size for region labels	10	
× location for region labels	(	•
y location for region labels	(	•
Labels	Only short labels	~
Use shadow effect	Use mass center	
Weight for distances to neighbouring labels	(	\$
Weight for overlap	(	•
ok		

## Showing river and cities

The NUTS2 map comprises geometry information about major rivers and cities above around 75.000 inhabitants, which can be added to the map:

Rivers Min width 4	
City labels Min. city size (1000000.0)	-

The label size for the rivers can be set as discussed above, however city labels are always shown in bold.



## Storing and re-loading your settings

Open the map option dialogue by pressing the map option button "

". Change the

settings according to your needs and then press the "store settings" button in the lower part of the dialogue. Choose a file name and a location. You may later use "load settings" to retrieve them again and apply them to another map.



## Copying the map to the clipboard or to a file on disk

In order to export the map to other applications, the easiest way is to use the clipboard, in order to do so, press the "copy to clipboard" 🗈 button. Afterwards, the current map can be imported into other applications as e.g. MS Word.

Another possibility is to store the current map in different format on disk, to do so, use the "export" button which will open a file dialog to choose the name of the file and select between different graphic formats. The "Options ..." button opens an additional dialogue

which is export type dependent, for PDF, to give an example, the paper size, orientation etc. can be changed. Generally, high quality are achieved if vector formats are used (PDF, SVG, EPS, EMF). These formats are however not supported by all applications, in opposite to e.g. bitmaps (BMP, which require a lot of disk space) and JPEG, which implies a loss of quality. It might therefore pay-off to try several formats for import into other applications. For MS Office users, the "Windows Enhanced Metafile" (EMF) format is interesting, as it allows changing later the graphics manually, e.g. by adding new text and changing colors.

The "Resolution factor" field allows to improve the quality of the saved file for non vector formats such as JPEG, by drawing the original graphic with more pixels which however drives up disk space.

#### Schemes shown with the mapping viewer

Technially, the mapping viewer colors polygons in appropriate columns and allows – as indicated above – to shrink the polygons according to the size of the numerical values. That allows applying the mapping viewer to graphically depict modeling result in a kind of flow chart. The example below depicts monetary flows in an economy simulated with a CGE shown as arrows and total account values as boxes.



Scenarios Tariffs

-10.63

Wolfgang Britz, Version March 2021

## Popup-menu in tables

A right click in the map opens a pop-up menu which gives quick access to some more often used functionalities:



Back to Table View Classify Classify current view by machine learning Customize .. Copy region info to clipboard Apply zom to all windows

- Back to Table View: Shows the data currently visualized as a table
- **Classify:** opens classification dialogue, see "Changing the classification and the legend"
- Classify current view by machine learning: see "Machine learning"
- **Customize:** opens the dialogue for general view options, see "Changing view options: fonts, number formatting and rounding, hiding empty cells, comparisons"
- **Copy region info to clipboard**: see the information on the info window in "Getting data for specific polygons". Its content is copied to clipboard.
- **Apply zoom to all windows**: If several map windows are opens, the currently zoom in regions will be (tried to) make also the zoom in all other map windows.

# Exporting the data to ArcGIS

The following section will briefly explain how to work with the data generated with the "Export to file" utility in DBF in ArcGIS. Under Layers, choose add Data

🕄 Untitled -	ArcMap - ArcInfo
Eile Edit Vie	ew Insert Selection Tools Window Help
🗅 🖻 🖪	😂   X 🖻 🛍 X   က က 🔶
	🔛 🖾 🖬 🖨 📑 100% 💆
	× :
💋 🥩 Laye	😎 Add Da <u>t</u> a
	💊 New <u>G</u> roup Layer
	🖹 Сору
	Paste Layer(s)

and in the case of the HSMUs, add the "HSMU\_EU27.shp" shapefile.

Add Data		×
Look in: 🔯 E:\capri1\GIS	<u> </u>	## ## #B
Name	Туре	
.svn	Folder	
🖾 capri.shp	Shapefile	
apri_MS.shp	Shapefile	
HSMU_EU27.shp	Shapefile	
1		
Name: HSMU_EU27.shp		Add
Show of type: Datasets and Layers (*.ly	r) 💌	Cancel

Then, choose add layers again, and add the dfb-file you have generated in the step explained above. You may also add the file with the meta data.

Add Data	X
Look in: 🔯 E:\CAPRI\gis	- <b>- - - -</b>
Name	Туре
SI_shrinked.shp	Shapefile
SK.csv	Text File
SK_shrinked.shp	Shapefile
💷 smu.dbf	dBASE Table
SMU.csv	Text File
🖾 solagro.shp	Shapefile
Std_cm01∨1.shp	Shapefile
Ⅲ test.dbf	dBASE Table
💷 test_meta.dbf	dBASE Table
UK.csv	Text File
Name: test dbf	Add
rescubi	
Show of type: Datasets and Layers (*.ly	r) 🔻 Cancel
, , ,	

Next, we need to connect the HSMU geometry with the newly loaded data, a process called "joining" in ArcGis. In the context of "HSMU\_EU27", choose "Join and Relates", then "Join …".

E:\capril\GIS	7 ArcToolbox	st Tools Tools
	Copy	ools
E:(CAPRI)gis	X <u>R</u> emove	ls
	III Open Attribute <u>T</u> able	rability Tools
	Joins and Relates	▶ <u>]</u> oin
	😪 Zoom To Layer	R <u>e</u> move Join(s)
	🛃 Zoom To Make Visible	<u>R</u> elate
	<u>V</u> isible Scale Range	▶ Remove Relate(s) ▶
	Us <u>e</u> Symbol Levels	r Tools cs Tools
	Selection	•
	Label Features	
	Convert Labels to Annotation	
	🍃 Convert Eeatures to Graphics	
	Convert Sym <u>b</u> ology to Representation	
	Data	•
	Save As La <u>v</u> er File	
<u> </u>	Properties	

That will open the join dialogue as shown below.

Join Data 🛛 🔀
Join lets you append additional data to this layer's attribute table so you can, for example, symbolize the layer's features using this data.
What do you want to join to this layer?
Join attributes from a table
1. Choose the field in this layer that the join will be based on:
FID GRIDCODE
2. HSMU
Show the attribute tables of layers in this list
3. Choose the field in the table to base the join on:
Advanced
About Joining Data OK Cancel

Make sure that "Join attributes from a table" is set in the first drop down box, and under 1., select HSMU, i.e. the filed in the HSMU\_27 geometry where the codes for the HSMU polygons are stored. Use the name of the exported dbf-table under 2., and select the field "Regions\_a" (the field name are restriced to 10 chars) under 3. Then press the button labeled "advanced", and chose the radiobutton "keep only matching records". If you are asked to build index, confirm.


If anything has worked well, you should now see the country or countries you had in the original map.

There is a trap, though. If you export several tables, or results for several scenarios, your table will normally have several fields used as a row header (e.g. year, scenario, activity). If that is the case, the join will not work properly as several rows for the same regions will be joined to the very same polygon. Unfortunately, ArcGIS will not warn you about that. *First* you have to execute a definition query in the table, while selecting the rows which you are later going to draw a map from.

In order to draw a thematic map now, it may be helpful to add the file with the meta data to the map, and to open the meta-data table with the help of its context menu. It will give us the long description and units belonging to the data fields in the exported data table.

▦	Attributes of test_meta							
	OID	Key	Name	Unit	LongText			
E	0	NMIN	No 5 (a), Mineral Fertilizer Consumption, Nitrogen	kg N/ha				
	1	PMIN	No 5 (b), Mineral Fertilizer Consumption, Phosphorous	kg N/ha				
	2	NMIN_SV/HE	No 5 (c), Mineral Nitrogen Application rate, Soft wheat	kg N/ha				
	3	PLAP	No 6, Consumption of Pesticides	Euro /ha		1		
	4	IRR	No 7(a), Irrigation, share	% irrigated				
	5	WAT_SURP	No 7(b), Irrigation, abstraction	l/m2				
	6	ELEC	No 8 (a), Energy, Electricity	Euro/ha				
	7	EGAS	No 8 (b), Energy, Gas	Euro/ha				
	8	EFUL	No 8 (c), Energy, Fuels	Euro/ha				
	9	LU	No 10 (a), Cropping/Livestock pattern, livestock density	Livestock units / ha UAA				
	10	RUMI	No 10 (b) Croppping/Livestock pattern, ruminants density	Livestock units / ha Fodder area				
	11	PP	No 10 (c), Cropping/Livestock pattern, non-ruminants density	Livestock units / ha UAA				
	12	ARAB	No 10 (d), Cropping/Livestock pattern, arable land density	%				
	13	GRAS	No 10 (d), Cropping/Livestock pattern, grass land density	%				
	14	PERM	No 10 (d), Cropping/Livestock pattern, permanent crops density(d)	%				
	15	INTE	No 12 (a), low-medium-high input farming	Index 0 - 2				
	16	H2865	BL21H2865					
	17	H2866	BL21H2866					
	18	H2867	BL21H2867					
	19	H2868	BL21H2868					
	20	H2869	BL21H2869					
	21	H2870	BL21H2870					
	22	H2871	BL21H2871					

Assuming we want to draw a map now with the ruminant stocking density, we find it in row 10 under the key "RUMI". In order to produce a map now, we have to open the context menu of "HSMU\_EU27", and choose properties, symbology, and choose "Quantities". Under values, choose "RUMI", the name before is the name of the DBF-file.

Layer Properties			? ×
General Source Select	tion Display S	ymbology   Fields   Definition Qu	uery Labels Joins & Relates
Show:	Draw quanti	ties using color to show valu	ies. Import
Categories	- Fields	Classification	
Quantities	Value:	none 💌	Manual
Graduated colors Graduated symbols	Normalization:	none HSMU_EU27.GRIDCODE	Classes: Classify
Dot density	Color Ramp:	HSMU_EU27.X test.OID	
Charts Multiple Attributes	Symbol Rar	test.GRAS test.LU test.PLAP test.ELEC test.EGAS test.EFUL test.NMIN test.ARAB test.PERM test.PP test.WAT_SURP test.NMIN_SWHE test.INTE	Label Advanced V
		0	K Cancel Apply

Afterwards, under classification, choose your preferred one. As there are many small polygons, the outline of the polygons should not be drawn. Therefore click on one of the colors, choose "Properties for all symbols ..." and under "Outline color" chose "No Color".

yer Proj	perties								?
General	Source	Selecti	on   Displa	ay Symbol	ogy Field	ds 🖡 Definition Que	ry Labels	Joins & Ro	elates
how:			Draw o	uantities u	sina col	or to show value	- 25		Import
eature	\$		– Fields –		<b>-</b>		– Classificat	ion	
ategor Iuantiti	les		Value:	Level I			Classificat	Ouantil	-
Grad	uated colo	nrs	value.	Ttest.r	NUMI			Guaria	
- Grad	uated sym	bols	Normaliz	ation:   none		<b>_</b>	Classes:	5 🗾	Classify
- Propo	ortional syr	mbols							
Dot d	lensity		Color Ran	np:		<b>•</b>			
harts:	An.:		Symbol	Range		L	.abel		
uiupie	Aundu	es		Clip Symbol	o ozoroc la	` ^	<b>~</b> 10000 - 0.8	73586	
				Pamp Colo	15 re		3587 - 1.4	71440	
				Properties	rs for Select	ed Symbol(s)	1441 - 2.0	68693	
				Properties	for All Sv	mbols	68694 - 3.3	54081	
	1 St	27					64082 - 147	67454617	6.000000
Æ	- <u>L</u>	/ (1		Reverse S	orting				
	アカフ			Combine C	lass(es)				
211	24	. 0	Shc -	Compilie C	102202		-		Advanced 👻
				Format La	oels			_	
				Edit Descri	ption				
						OK		Cancel	Apply
nbol Sele	ector						?)	<	
togon	All				-	- Preview		? ×	
tegoiy.	I AII								
<b>C</b>		Dhar		C					
Gree	'n	Blue		Sun					
	_		_			Options		μш	
						Fill Color:	-		
Hollo	W	Lake		Rose		Outline Width:	0.40 🖃		
			_				0.40		
						Outline Color:	-		
								No Co	lor
Beig	e	Yellov	v	Olive					
						_			
						Propertie			
Gree	n.	Jade		Blue		<u>M</u> ore Syn	nt 🔜 🔳 📕		
0,00		5000		E.42					
						Save			
					-	ОК			
								More Cel	
								more Col	Urs

Afterwards, if everything went well, you should see your map.



## **Machine learning**

### Motivation

A serious challenge for large-scale economic models is the dimensionality of the results generated by model runs. These reflect the high level of dis-aggregation in different dimensions and the many aspects dealt with in these tools, such as relating to economic, social and environmental indicators. A single simulation run e.g. with CAPRI based on the farm type modules produces over 20 Mio non-zeros. Clearly, any of these numbers is generated by a deterministic computer based model and should hence be a deterministic outcome depending on the input and the code used. Specifically, the relation between the input and any single number outputted is determined by the model structure and parameterization, and pre-and post-processing code. It must hence be possible to track any change quantitatively back to the shock analyzed.

But that rather theoretical point of view has very little to do with the task at hand when one has to distill an analysis from a set of model outcomes. The questions here are: what are the most important results, i.e. salient to the questions underlying the analysis and large enough to matter, and how can they be explained? For the client, the story behind the results is often at least equally important as the results themselves. If the story is well told, the "black box" character of the tool is removed and its usefulness in depicting major cause-effect relations becomes evident. Telling a good and right story requires however often quite some time in analyzing results in a systematic way.

The user will hence have to decide for which items of the huge data set a thorough analysis of underlying drivers is advisable. Limited time and human resources will set tight limits to the extent of such systematic analysis. Typically, in any report, only a few dozen key results (perhaps complemented with a few maps showing several hundredths numbers) will be presented. But these key results, such as changes in aggregate welfare, farm income, GHG emissions or the nitrogen balance are calculated from thousands of simulated items. How can we discover "the story behind the results", i.e. which regions, activities, price or policy changes etc. are most important for the aggregate changes communicated?

The exploitation tools of GGIG with a flexible on-the-fly approach to produce tables, graphs and maps had been an important step to improve the efficiency in exploiting and analyzing results. But in parallel, tools such as CAPRI has grown in scope and scale. It might be the time now to consider new approaches to analyze model outcomes.

Wikipedia gives the following definition: "Machine learning, a branch of artificial intelligence, is a scientific discipline concerned with the design and development of algorithms that allow computers to evolve behaviors based on empirical data, such as from sensor data or databases. Machine Learning is concerned with the development of algorithms allowing the machine to learn via inductive inference based on observation data that represent incomplete information about statistical phenomenon. Classification which is also referred to as pattern recognition, is a important task in Machine Learning, by which machines "learn" to automatically recognize complex pattern, to distinguish between exemplars based on their different patterns, and to make intelligent decisions."

That is naturally a very general description. Machine learning has been widely in a wide range of application fields. A typical example is the analysis of which clients of a bank has been given credits. We have many observations with "credit granted" or "credit refused", and probably a longer list of attributes of the clients (age, sex, income, amount of the credit asked for, time since being a customer with the bank, past bookings ... ). Machine learning could be applied to define a set of rules which based on past decisions predict if a credit would be granted for a new application or not. Machine learning will in many cases also be able to tell something about the possible error range linked with the decision. That could e.g. allow the banks to make fast decisions in many cases, and spend more time on the tricky ones. The book by Witten et.al. 2011 gives many such real world examples. It might be interesting to note that one of the often cited applications of the WEKA package used in CAPRI refers to agriculture (Queen et.al. 2005), as the WEKA authors write: "New Zealand has several

research centres dedicated to agriculture and horticulture, which provided the original impetus for our work, and many of our early applications." The applications deal mostly with micro level data. A recent application to price agricultural forecasting is found in Ticlavilca et.al. 2010.

With the approach used by machine learning, we can e.g. interpret the income changes in each farm type in a simulation compared to the baseline as an outcome we want to predict, and the production program of each farm type and changes in prices and premiums as the attributes used to explain that outcome. Some farm types might exhibit very large income changes,



other little ones. What are common characteristics of the one and the other group? Machine learning might then come up with a "pattern" (e.g. based on a regression model) which determines the most important attributes impacting income changes in a given simulation. Machine learning has thus a lot of similarities with statistics – indeed many methods can also be found in

statistical packages - but the focus to decide upon which attributes and relations matters is shifted to a certain extent from the human being to the computer. And, the tool box used in machine learning differs to a certain degree from classical statistics. And, not of least, many of the algorithms had also been developed keeping computing time in mind.

## Implementation in GGIG

The implementation in GGIG is based on integrating the WEKA machine learning library (Witten et.al. 2011) into the GGIG exploitations tool, which is possible thanks to the GNU license which includes full access to the underlying Java source code. WEKA is also integrated into other well known packages such as RapidMiner. Relatively few code changes were necessary in the GGIG Java code to pass data from the tables and maps shown in the GUI to the WEKA library (see below). That is done automatically in the background with the aim to reduce user input in the process.

As a consequence, a very powerful set of filtering and classification as well as related visualization tools from machine learning can be applied to the result sets from CAPRI inside the existing exploitation tools.

The current implementation is based on the interaction of two views:

- A **map** or a table using classification colors it defines the class attribute (=dependent variable) of the data to classify. For classification algorithms which require nominal values, the assigned class from the classification defining the color scheme is used.
- A **table** with the "explanatory" attributes.

Both tables must be, as conventionally in the exploitation tools, the observations in the rows. For maps, each map carries the data for a region. But one might also work with tables where the observations are not strictly geo-referenced entities such as farm types.

The GUI will automatically send new data to the WEKA GUI if either the map (or the table using classification colors) or the table providing the explanatory attributes is updated by a user action. The basic data flow is shown in the graphic below.



#### Interaction between the GGIG GUI and WEKA

Let's construct an example based on CAPRI: we want to check if the income change in cereals in a simulation depends on the crop shares of cereals and the yields. In order to do so, we first render our map as usual (table "Farm details, mapping view", use the option dialogue to show percentage changes against the baseline):



The regions shown are our instances and the value plotted for a region defines the class attribute we want to analyze. Any one instance consists of a vector of attributes of which one is the "class value", i.e. the value to classify, which can be numeric or nominal. The other attributes are used for classification or clustering and stem from a second table (see below). Classification methods which use nominal values can also be used. In that case, the class chosen for the region, as seen from the color in which is drawn, defines the class attributes. In our example above, each region would fall into one of five classes.

Next, we open a second table with the data we want to use as explanatory attributes. The latest trunk comprises the table "Supply details, cluster view" which comprises promising attributes which are possible candidates to explain many changes in a simulation (for all activity aggregates: crop shares/stocking densities, revenues, income, yields).

2 Sunkedabile clusteriaer[0]									
Years	[4]		s	cenarios					
2020				MTR_RD				-	
	Cereals Revenues [Euro/ha or head]	Cereals Income [Euro/ha or head]	Cereals Yield [kg or 1/1000 head/ha or head]	Cereals Crop share/Animal density [% or 0.01 animals/ha]	Oilseeds Revenues [Euro/ha or head]	Oilseeds Income [Euro/ha or head]	Oilseeds Yield [kg or 1/1000 head/ha or head]	Oilseeds Crop share/Animal density [% or 0.01 animals/ha]	Other arable Revenues [Euro/ha or l
European Union 27	818.70	468.75	5524.26	30.49	893.92	617.03	2806.03	4.90	i
European Union 25	839.26	482.20	5758.41	30.13	994.25	566.36	3105.66	4.28	,
European Union 15	952.81	513.61	6318.16	25.80	1065.26	566.93	3388.98	3.45	1
European Union 12	609.81	400.97	4278.35	42.65	722.94	475.00	2187.93	8.54	ı
European Union 10	578.65	419.93	4513.20	48.82	861.86	585.72	2563.58	7.66	,
Belgium	1177.10	661.93	8640.45	24.77	1409.88	778.13	4336.86	3.76	i .
Benmark	1013.76	352.26	6855.81	53.84	1249.95	497.01	3778.61	3.26	
Germany	1059.35	441.45	7527.86	39.25	1314.72	638.80	4251.82	7.29	1
Austria	889.55	497.57	6973.35	21.66	947.31	675.02	2404.66	3.17	
Netherlands	1266.88	1 711.29	8827.19	12.49	868.24	764.83	3095.68	0.52	1
France	1085.12	438.20	7479.60	29.96	1014.91	456.67	3405.41	6.27	
Portugal	746.24	356.13	3898.34	6.25	191.71	144.97	497.99	0.60	
Spain	593.53	632.23	3498.86	20.42	433.16	473.81	1072.65	1.29	
Greece	794.69	891.07	4133.29	18.20	608.18	776.33	1655.66	0.84	i
Italy	1053.24	706.05	6754.93	28.54	874.83	542.60	2871.15	1.68	,

In order to start the clustering/classification, we click in the table to open its popup-men and then select "Classification":

🔬 CAPRI [t:\britz\capri\gams]	🕼 CAPRI (t\britz\capr\gams)												
View Handling Windows	View Handling Windows												
🛃 Supply details, cluster view	🔝 Supply details, clusterview (0)									6 🗙			
Years			S	cenarios							Ba	V	New type
2020				MTR_RD				-					Table
	Cereals Revenues [Euro/ha or head]	Cereals Income [Euro/ha or head]	Cereals Yield [kg or 1/1000 head/ha or head]	Cereals Crop share/Animal density [% or 0.01 animals/ha]	Oilseeds Revenues [Euro/ha or head]	Oilseeds Income [Euro/ha or head]	Oilseeds Yield [kg or 1/1000 head/ha or head]	Oilseeds Crop share/Animal density [% or 0.01 animals/ha]	Other arable crops Revenues [Euro/ha or head]	Other arable crops Income [Euro/ha or head]	Other arable crops Yield [kg or 1/1000 head/ha or head]	Other arable crops Crop share/Animal density [% or 0.01 animals/ha]	Veget Perm Rever (Euro.
European Union 27	818.76	468.75	5524.2	30.49	893.92	517.0	3 2806.03	4.90	3274.41	1309.45	22201.53	3.96	5 _
European Union 25	839.28	482.20	5758.4	30.13	994.25	566.3	3105.66	4.20	3261.98	1242.50	23464.87	4.03	з Е
European Union 15	952.81	513.61	6318.10	5 25.80	1065.26	566.9	3 3388.98	3.45	3529.35	1301.68	23704.28	4.25	5
European Union 12	609.81	400.97	4278.3	5 42.65	722.94	475.0	2187.93	8.54	2286.71	1232.57	16985.51	3.18	8
European Union 10	578.65	419.93	4513.2	9 48.02	861.86	585.7	2 2563.58	7.56	1688.16	848.92	22102.42	3.10	a
Belgium	1177.10	561.93	8640.4	i 24.77	1409.88	778.43	4336.85	3.76	4412.24	1833.10	43975.38	12.23	3
Denmark	1013.76	352.26	6855.0	53.84	1249.95	497.0	3778.61	3.26	2916.66	757.41	22763.14	6.34	1
Germany	889.55	441.40	6973 3	5 24.66	947 34	636.8	2484.66	1.2	2477.94	729.63	44224.23	4.63	5
Austria	1266.88	711.29	8827.1	Reload	868.24	764.8	3095.68	0.53	9119.67	3764.04	48370.33	13.47	,
France	1085.12	438.20	7479.61	Export	Data 1014.91	456.6	7 3405.41	6.21	3865.74	1820.01	42422.31	3.03	3
Portugal	746.24	356.13	3898.3	Pivotin Custor	nize Table 191.71	144.9	7 497.99	0.60	53672.36	32871.44	41242.75	0.22	2
Spain	693.63	632.23	3498.84	Statisti	cs433.16	473.8	1072.65	1.25	2071.26	-698.01	4844.85	6.19	9
Greece	794.65	891.07	4133.2	Classify	y ▶ Ci Ci	assify numeric 76.3	3 1655.66	0.04	4215.46	3829.66	2709.12	8.62	2
Raly	1053.24	706.05	5754.9	Table \	/iew	o not classify 42.6	2871.15	1.68	4297.25	2961.06	11189.96	2.55	6
Ireland	1082.81	676.66	8583.7	0.10	1320.34	889.3	4 3681.85	0.02	7271.07	-1208.30	16318.50	0.73	3
Finland	477.95	393.91	3614.8	47.08	417.17	612.7	1258.13	3.26	2517.05	965.92	23951.47	2.04	
Sweden	751.55	251.31	5491.8	5 47.47	1285.84	381.4	2741.29	1.62	2865.73	2695.7	30978.91	3.04	2
United Kingdom	730.23	604.25	5049.9	45.72	869.85	663.6	2396.48	13.74	1956.81	726.90	20829.30	5.08	8
Estonia	462.43	364.06	4001.3	3 44.74	775.63	679.0	2255.56	10.52	1419.67	706.33	8587.71	0.81	1
	1		6674 M		040 Z /	ena 41	9200 40	45 74	2400 70	4996 70	24376 34		• •
CAPRI GUI Version 4.0 , Septe	mber 2011 II	ni file : capri.ini		User name : Wol	fgang Britz	User typ	e : administrator						
🚱 🥝 🐁				<u></u>							DE ,	· P 😰 🕪 🕺	(1:59 09.2011

We clicking one of the option if we can then decide to:

- use *numerical classification* methods such as different regression methods. The observations in the map define the dependent variable.
- Use the class assigned by the maps input into *nominal classification*.
- To switch classification off.

A new window will be opened which shows the WEKA GUI (see below).

## The WEKA GUI

The classification is based on the complete functionality of the WEKA GUI regarding attribute selection/visualization, filtering and classification, see

<u>http://www.cs.waikato.ac.nz/~ml/index.html</u>. There are very good manuals available from the site (the latest user manual is also available from <u>http://www.capri-</u>

<u>model.org/docs/WekaManual-3-6-5.pdf</u>), so that only a few major tips are given below for fast start.

The tabs "Classify", "Cluster", "Filter" and "View and select" allow the user to access specific part of the WEKA functionality. The result set from the current classification run can be shown in the lower left panel (result list). For each result set, a popup menu opens options, e.g. to show a graph with the prediction errors.

### Classification

assify Cluster Filter View and Selec	t		
lassifier			
Choose LinearRegression -5 0	-R 1.0E-8		
est options	Classifier output		
<ul> <li>Use training set</li> </ul>	0.00/3 * Fodder activities Revenues +		
	-0.0085 * Fodder activities Income +		
Supplied test sec Set	-U.164 * Set aside and fallow land Revenues +		
Cross-validation Folds 10	0.0042 * Set aside and fallow land income +		
Percentage split % 66	-0.0048 % All cattle activities Income +		
	-0.0268 * All cattle activities Yield +		
More options	0.005 * Beef meat activities Revenues +		
	0.0026 * Beef meat activities Income +		
lum) Class-numeric	<ul> <li>0.0207 * Beef meat activities Yield +</li> </ul>		
2.6	-0.0019 * Other animals Revenues +		
Start Stop	0.0031 * Other animals Income +		
esult list (right-click for options)	-9.7609		
2:02:01 - rules.ZeroR			
2:02:28 - functions.LinearRegression	Time taken to build model: 0.04seconds		
	=== Evaluation on training set ===		
	Sunnary		
	Correlation coefficient 0.807		
	Mean absolute error 2.4774		
	Root mean squared error 3.369		
	Relative absolute error 61.4055 %		
	Root relative squared error 59.051 %		
	Total Number of Instances 269		
	Ignored Class Unknown Instances 31		

- The "choose" button will give access to a wide range of different classifiers, many of which have additionally options which can be edited by users. A multiple linear regression using the Akaide criterion for model selection is used as the default, assuming that most people will start with using numerical values as class attributes. Please not that switching between nominal and numerical class attributes might trigger error messages if the currently selected classifier cannot handle the newly selected class attribute type.
- It is recommended for our purposes to use under "Test options" "Use training set" (the default in our implementation) as we are typically not interested in an out-of-sample test of the prediction quality.

- The actual classification can be started with the "start" button. If the data in the background are updated, the actually chosen classifier with the chosen options will be started on the new data set automatically. In absence of errors the "Classifier output" on the RHS will hence typically show results based on the latest selected data.
- The results can be visualized by clicking with the mouse on an item in the result list, the last on in the list always being the newest. If one has tried several classifiers, the old results remain available. However, if the data in the background change, the old results are automatically removed.

The reader should note that all the functionality described is from the standard WEKA GUI so that the user manual from WEKA can be used for further information.

PS: The cluster panel is not described, it works quite similar. Note however that filters are not applied to the cluster (see below).

### Filtering



The filter panel allows running different types of filters which remove attributes, in many cases reflecting the correlation between attributes. In order to use the result from the filter run, click on the result set in the result list and chose "Use output for classification":

Weka Explorer GUI Supply details, mai	pping view [0] : all.Cereals.Income.2020.MTR_RDOuantile : 20:26:35 -			
Classifie Cluster Eliter, View and Select				
Attribute Evaluator				
Character Character				
Choose Crssubsetteval				
Search Method				
Choose GeneticSearch -Z 20 -G 2	20 -C 0.6 -M 0.033 -R 20 -S 1			
Attribute Selection Mode	Attribute selection output			
Ise full training set	0.33361 0 16307101317136061663135			
	0.48977 0.44828 1 2 5 14 17 19 21 22 24 30 31 36 37 39			
Cross-validation Folds 10	0.52524 0.56472 2 5 6 8 14 17 20 21 22 24 30 31 33 34 37			
Seed 1	0.53106 0.58384 2 5 6 8 14 17 19 20 22 27 31 36 37 39 40			
(Num) Class-numeric 🗸	Attribute Subset Evaluator (supervised, Class (numeric): 46 Class-numeric): CFS Subset Evaluator			
	Including locally predictive attributes			
Start Stop				
Result list (right-click for options)	Selected attributes: 1,2,5,6,7,10,17,19,20,21,22,27,31,37,39 : 15			
20:26:51 - GeneticSearch + 61 C 1 - 15	Comcals Revenues			
Use ou	tput for classification als Income			
Do not	clongeruse output for classification als Production per UAAR			
	teds Revenues			
View in	main window teds Income			
View in	separate window eds Production per UAAR			
Save re	sult buffer			
Delete	result huffer and Dermanent groups Broduction per UhiD			
	r activities Revenues			
Visuali	ze reduced data er activities Income			
	Set aside and fallow land Income			
	All cattle activities Revenues			
	Beef meat activities Income			
	Beef meat activities Crop share/Animal density			
	K			

The last selected filter will be automatically restarted if a new data set is implicitly loaded (change of the map or of the data in the cluster table with the explanatory results). In order to switch off the use of the filter, select "Do not longer use output for classification"

#### Attribute viewing and selection

The last panel available is especially interesting to quickly analyze statistics of the underlying data:

🕼 Weka Explorer GUI Supply details, mapping view [1] Quantile : 20:46:17 -							
Classify Filter View and Select							
Use modified data for classification							
Elter							
Chorse None		Apply					
Gurrent relation	Selected attribute						
Relation: Supply details, mapping view [1] Quantile : 20:46:17 - Instances: 300 Attributes: 5	Name: Revenues Missing: 0 (0%) Distinct: 300	Type: Numeric Unique: 300 (100%)					
Attributes	Statistic Valu	Je					
All None Invert Pattern	Minimum 2.07	76					
	Maxmum 2190 Mean 877	0.78					
No. Name	StdDev 352	.113					
	Class: Class-numeric (Num)	- Visualize Al					
Remove							

The reader can manually remove attributes and the reduced set of attributes will then passed to the filter and classifier. However, the attribute selection is not maintained when new data are loaded. The "Visualize All" button produces graphs of all current attributes:



### Summary

The integration of algorithms from machine learning based on the WEKA library and GUI offers new possibilities to systematic analysis of result sets. Thanks to the open source policy of WEKA, it was possible to integrate these powerful tools transparently in the CAPRI GUI. Depending on the experiences made over the next months, further links might be included (e.g. rending clusters in maps).

#### References

Ian H. Witten, Eibe Frank, Mark A. Hall (2011). Data Mining Practical Machine Learning Tools and Techniqes. Third edition. Elsevier, Amsterdam. 630 pages

Remco R. Bouckaert, Eibe Frank, Mark Hall, Richard Kirkby, Peter Reutemann, Alex Seewald, David Scuse (2011). WEKA Manual for Version 3-6-5. June 28, 2011, University of Waikato, Hamilton, New Zealand.

Robert J. McQueen, Stephen R. Garner, Craig G. Nevill-Manning, Ian H. Witten (1995). Applying machine learning to agricultural data, Computers and Electronics in Agriculture, Volume 12, Issue 4, June 1995, Pages 275-293, ISSN 0168-1699 (http://www.sciencedirect.com/science/article/pii/0168169995986019)

Ticlavilca, A. M., Dillon M. Feuz and Mac McKee. 2010. "Forecasting Agricultural Commodity Prices Using Multivariate Bayesian Machine Learning Regression." Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. [http://www.farmdoc.illinois.edu/nccc134].

# Scenario editor

### Concept of the scenario editor

The scenario editor is an optional tool to be embedded in a GGIG user interface which supports the user in setting up run specific include files where the content is *not* stemming from GUI controls. That parallel way to define run specific input is typically necessary for more complex tools where e.g. policy scenarios are defined in GAMS code and not via controls on the interface.

The scenario editor is a "predefined" task which must be named "Define scenario", e.g.

```
<task>
<name>Define scenario</name>
<userLevls>runner,Administrator,developper,debugger</userLevls>
</task>
```

The editor allows defining separate directories for so-called pre-edited "snippets" which can be edited and combined by the user to scenario files and the scenarios files generated by the user. These files are saved to a separate directory tree and can later be selected and included in a GAMS application.

The related file dialogs let the user now only navigate under these directories and use filters for the "gms" extensions.

The two relevant fields in the XML are:

```
<scenarioDir><attr>scen</attr></scenarioDir><userScenarioDir><userScenarioDir><attr>userScens</attr></userScenarioDir></userScens</attr></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></userScenarioDir></
```

Where "gamsdir" is added in front if the directory is not found.

#### Using the scenario editor

The screen shot below shows an example from GTAP.



Choosing the scenario editor adds the panel with GUI elements shown above. That panel consist of two major panes:

- 1. A **top pane** where the user can enter a name for his new scenario (required), and a description text (optional).
- 2. A bottom pane where the user can define the base scenario to start with and the snippets to add. The available snippets and their structure are shown on the left hand side in an expandable tree which shows the sub-directories found (typically under "gams\scen"), with the exclusion of a sub-directory called "baseScenarios" and the ".svn" directories. Empty directories are not shown. The user may select any number of snippets, even several ones from the same sub-directory. Double-clicking on one of the snippets shows the content of the file on the right hand side, so that the user can inspect the code as seen below in more detail. GAMS keywords are shown in red, comments in yellow and strings in green. She can also edit the file changes are shown in blue. Once changes had been saved, the tree shows "(user modified)" behind the modified snipped. The user can also remove these changes again from snippets. In the upper left part of the button are controls which allow for search for snippets as detailed below.

### Working with the policy editor

The user can select one or several "snippets" in the tree plus add a basis scenario. The content of these files are stored on demand to a new file. The steps are as follows:

1. The **selection of the snippets** is done in the tree. An example is shown below where two snippets are selected.



In that case where the snippets had not been modified by the user, the generated file comprises only includes:

*******	***************************************
*	
* author	: undefined
* called by	COMgms
* date	: 23-06-2016 07:14:50
* purpose	: Scenario definition
*	
*******	*******************************
	NDEC
	NDL3
* Baseline scer	ano
\$include\aan	ns\scen\base_scenarios\noShock.gms
*	
* Category : F	Policy Shocks
\$include\gan	ns\scen\Policy_Shocks\Direct_taxes.gms
*	
* Category : F	Policy Shocks
\$include\gan	ns\scen\Policy_Shocks\Consumption_taxes.gms

2. Editing of the snippets is done on the editor to the right. In the example below, the numerical value was changed to 0.4 as seen from the underlined character. Comments are shown in green, GAMS keywords in red, pre-processor commands in orange and strings in grey.



Once a snippet is edited, the "Save changes" and "Save to disk" (the latter only for administrator") and the "Reset changes" button become visible. If the user press "Save changes", a temporary copy of the snipped with her changes is stored. That can be seen in the tree by the add "(user modified)" behind the snippets' name.

Scenario categories	
🗉 🔄 base scenarios	tifthani "04 inclued dot4""deal"
🖶 🔄 End and tp shocks	
Parameters	
Policy Shocks	
Consumption taxes	* define which regions and factors to shock
Direct taxes ( user modilied)	define which regions and factors to since
full ib	the official and the kill scattering
Tariffs	alow to use that coinst duction in several nea
III- Test Shocks	
user scenarios	
• xxx	\$onmulti
	set (Shocked(f):
	Softmulti
	option kill=rShocked;
	option kill=fShocked;
	rShocked(r) = YES;
	fShocked(f) = YES;
	parameter p_taxShockDirect(r,t);
	Field, we could e.g. also lister table, read non
	n_taxShockDirect(rShocked (Shocked) = 0.4
	\$else
	* r = importer regions, f=factor
	kappaf. <mark>IX</mark> (r,f,tSim) \$ kappaf. <b>I</b> (r,f,tSim) = p_taxShockDirect(r,f) * kappaf. <b>I</b> (r,f,tSim);
	Sendit

In case that scenario files are generated from user modified snippets, instead of an include, the modified file content is stored to disk (the screen shot below only shows the top part of the content of the edited file):

******	*****	*******	******	****
* author * called by * date * purpose	: undefined : COMgms : 23-06-2016 07:21:47 : Scenario definition			
* * * * * \$ setglobal SCE	**************************************	******	******	****
* Baseline scer * \$include\gam *	nario ns\scen\base_scenarios\no:	Shock.gms		
* * * \$ontext	*****	******	******	******
GAMS file : D @purpose : @author : N	rect_taxes Cut direct taxes on factor : Wolfgang Britz	income		
@date : 1 @since : @refDoc : @seeAlso : r @calledBy : (	nodel.gms, comgms			
siftheni "%incli *	**************************************	***************	*******	*****

Storing the scenario then generates a file as shown below, user name, the reference to CAPMOD.GMS and the date and time are automatically added by the GUI. The files will be added to the files stored in "gams\pol\_input" (or an alternative directory defined in the GGIG definition file)

3. **Generating a new scenario file**. First, a scenario name should be added in the upper most text field after "Enter scenario name". Optionally, a long text description of the scenario can be entered in the "Enter scenario description".

Scenario description	Enter scenario nam	ne DirTaxesMod	
Enter scenario description	lified direct tax scenario with a reduction level of 40%		^ v
-	1		
Next, the store s	cenario button is press:	Store scenario	which opens a file save

dialog. The dialog allows also to generate new directories and to store the file into such as new directory:

🐡 Speichern		×
Speichern in: 📙 user	_scenarios ~	🦻 📂 🖽 -
<ul> <li>↓ test</li> <li>▲ biomass.gms</li> <li>▲ capTax80.gms</li> <li>▲ capTax80Etaf.gms</li> <li>▲ etaf.gms</li> <li>▲ eth_notariffs.gms</li> <li>▲ fta_gha_nga.gms</li> </ul>	<ul> <li>Kappa.gms</li> <li>Xxx1.gms</li> <li>OilGasShockUS.gms</li> <li>yyy.gms</li> <li>permistEU.gms</li> <li>permits.gms</li> <li>test.gms</li> <li>TFP_20.gms</li> <li>xxx.gms</li> </ul>	
Dateiname: DirTaxes	Mod.gms	Speichern
Dateityp: GAMS file	2S	~ Abbrechen

If the file already exists, a warning message appears:

🐡 Speichern	×
Speichern in:	user_scenarios v 🤌 📂 🖽 🗸
test biomass.g capTax80 capTax80 DirTaxesM etaf.gms eth_notar	<ul> <li>▲ fta_oha_noa.oms ▲ xxx.oms</li> <li>Ims ▲ Kap Existing file ×</li> <li>OilG</li> <li>OEtaf.gms ▲ perr</li> <li>A oliG</li> <li>The file exists, overwrite?</li> <li>Ind.gms ▲ perr</li> <li>Ind.gms ▲ TFP</li> </ul>
Dateiname:	DirTaxesMod.gms Speichern
Dateityp:	GAMS files ~ Abbrechen

Next, the tag editor is opened (see next section) which alows to add meta-information to the file. Afterwards, the file is added to the "user scenarios" and the editor updated.

#### **Tag editor**

So far, when the user stored a scenario file by pressing the "Store scenario" button, three default tags were generated automatically in the new file:

*:	*******	***************************************
*		
*	author	: undefined
*	called by	: COMgms
*	date	: 21-06-2016 10:11:13
*	purpose	: Scenario definition
*		
*:	*******	***************************************
*		

Where "purpose" was always set to "Scenario definition". The other fields use information from the setting dialog, the current time and the GAMS file registered with the editor.

The improved version of the editor now shows potentially four dialogs:

- 1. A warning dialog of the scenario description field is empty.
- 2. A file selection dialog.
- 3. A warning dialog if the selected file is already existing.
- 4. A tag editor.

The later allows the user to edit all tags which are currently found on any file, with the exemption of the "author"," called by", "date" and "purpose" tags which are automatically generated :

page -	×
calledby	
purpose	
refdoc	
seealso	
since	
0	k

#### Search functionality

The policy editor allows searching for files in all tags, i.e. the header lines of the files, or alternatively, to perform a full text search in the non-header part of the files. A tag is a line in the file header where the tag's name on the right is separate from the tag's content on the left:



The user can select any tag name and either user a free search field or select from a list of found contents (see below). The search field populates a selection of files where matching content is found; at the same time, it updates the list with the contents of the tab.

The tags are read dynamically from the files, i.e. the development team can decide for themselves which tags they use (e.g. "policy context", "key words", "project"). However, the "purpose" tag is special as it defines a tooltip for each file in the tree. The first field in the list of tag names is empty:



and has a special meaning. It allows for a full search text in the main body of the file, e.g. to search for specific GAMS symbols.

The different fields and their functionality is shown below:



Clicking on a file will select that file in the tree, double clicking on file in the tree loads the file in the editor:



The search functionality is most useful all files are properly "tagged", i.e. if there file header comprises pre-defined search fields, so-called tags, with brief and clear descriptions. Potential useful tags for IFMCAP are for instance "Key words", "Policy context" and "Project". The aim of these tags is to let a user quickly find files which comprise code relevant for the scenario she intends to implement. To ease that process, a tag editor is not opened when a new user scenario is generated as discussed next.

#### Update possibility for administrator

So far, the editor did not allow editing the pre-defined "snippets". The edits of the user where only temporarily kept in memory and only saved to the (newly) generated scenario files. Now, in administrator mode, there is a new button "Save to disk" to save the changes of the current snipped either to the current file or to a copy. That allows, for example, updating or adding tags or generating variants of the files without the need to use an external editor.

It should be noted here that updating the processing content of existing snippets can potentially lead to surprises for users if there scenario files use \$include statement to refer to the snippet. The user has decided to include these snippets based on the content seen before the potential changes.

parameter p_taxShockTariffs(r,i,rp) p_taxShockExpSubs(r,i,rp);
<ul> <li>* here, we could e.g. also insert a table, read from</li> <li>* EXCEL or from a GDX container</li> </ul>
p_taxShockTariffs(rShocked,iShocked,rpShocked) = 0.5; p_taxShockExpSubs(rShocked,iShocked,rpShocked) = 0.4;
<u>\$else</u>
<ul> <li>r = importer regions, i=products, rp= exporters, tsim= simulation time point</li> <li>imptx: Bilateral import taxes</li> </ul>
imptx. <mark>fx</mark> (r,i,rp,tsim) \$ imptx. <b>(</b> (r,i,rp,tsim) = p_taxShockTariffs(r,i,rp) * imptx. <b>(</b> (r,i,rp,tsim); exptx. <b>fx</b> (r,i,rp,tsim) \$ (exptx. <b>(</b> (r,i,rp,tsim) <b>0</b> ) = p_taxShockExpSubs(r,i,rp) * exptx.((r,i,rp,tsim);
Şendif <
Store scenarioSave changesSave to diskReset changes

## Meta data handling

### Why meta data?

Meta data are data about data. In many GAMS projects, it is impossible or cumbersome to tell solely from a look at the numerical values exactly based on which shocks and settings results of a model run had been generated, especially if results are stored separately from the listing file, e.g. in GDX containers or EXCEL sheets. But even with a listing file, important settings such the state of environment variables e.g. used in "\$if" statements might not be reported. And typically, listing files are regularly overwritten. Later on, result users are often left guessing what exactly the settings underlying the run might have been. In order to overcome that problem, the GGIG, drawing on former <u>CAPRI GUI concepts</u>, passes all interface settings, plus the user name and the current time, forward to GAMS in one SET called META. Additionally, the names of all controls are passed in a second set named allMetItmes.

A correctly defined interface with GGIG should allow to steer *all* run specific settings. If that is the case, the meta data generated by GGIG will provide an exact and sufficient definition of all run specific inputs, ensuring that all relevant meta data about a run are stored along with quantitative results in the same GDX file. Accordingly, GDX files shipped to other desks or committed e.g. to a SVN server still carry all necessary information to identify exactly the run.

## Technical concept

The meta handling is straight forward. The state of the different control is reported as elements of the set META and related long text descriptions report the state of the control as shown below from an example application:

÷.	SET S_META /	
1	'Simulation'.	'Workstep' 'Simulation'
e.	'Simulation'.	'Task' 'Simulation'
i.	'Simulation'.	'Date and time' '2016-06-23 15:49:21'
ŧ.	'Simulation'.	'User' 'undefined'
÷.	'Simulation'.	'Gams version' 'GANS.%system.GamsRelease%'
i.	'Simulation'.	'Standard GTAP model' 'true [default]'
t.	'Simulation'.	'Input file (*.qdx)' '10x10 example [default]'
e.	'Simulation'.	'Scenario file' 'xxx'
i.	'Simulation'.	'Dynamics' 'Comparative static [default]'
ŧ.	'Simulation'.	'Number of years' '10 [default]'
e.	'Simulation'.	'Regional coverage' 'Global model [default]'
1	'Simulation'.	'Region in model' 'Oceania'
t.	'Simulation'.	'Non default parameters' 'false [default]'
1	'Simulation'.	'Parameter file' 'CET'
1	'Simulation'.	'Residual region' 'Oceania'
t.	'Simulation'.	'Numeraire products' 'GrainsCrops-c,MeatLstk-c,Extraction-c,ProcFood-c,TextWapp-c,LightMnfc-c,HeavyMnfc-c,Util Cons-c,TransComm-c,OthServices-c'
1	'Simulation'.	'Numeraire regions' 'Oceania,EastAsia,SEAsia,SouthAsia,NAmerica,LatinAmer,EU_25,MENA,SSA,RestofWorld'
÷.	'Simulation'.	'hide_Mobile factors' 'Land,UnSkLab,SkLab,Capital,NatRes'
t.	'Simulation'.	'Foreign savings' 'Global equal returns to capital [default]'
1	'Simulation'.	'Current account balance' 'Fixed foreign savings [default]'
i.	'Simulation'.	'Government' 'Tax income [default]'
1	'Simulation'.	'Final consumption' 'Spending [default]'
1	'Simulation'.	'Exchange rate' 'Monetary Union [default]'
i.	'Simulation'.	'Factors with factors tax udates' 'Land,UnSkLab,SkLab,Capital,NatRes'
1	'Simulation'.	'Products with consumption udates' 'GrainsCrops-c,HeatLstk-c,Extraction-c,ProcFood-c,TextWapp-c,LightMnfc-c,HeavyHnfc-c,Util_Cons-c,TransComm-c,OthServices-c
1	'Simulation'.	'Output' 'GUI,GDX [default]'
τ.	'Simulation'.	'Autout for flow schemer' 'true [default]'

Having the current task as the first dimension of the set "S\_META" opens the possibility to add meta data from results of other tasks used in a run and add them to the meta data in a very simple way:

```
*
*
* --- load meta information from data set
* (date of production, type of filtering, tolerances)
*
$
iftheni defined s_meta
execute_loadpoint "%datDir%\%dataset%.gdx" s_meta;
$endif
```

That sets can then be stored, with one GAMS statement as shown below, in a GDX file along with the results:

#### execute\_unload "%scenario\_description%.gdx" META,RESULT;

The user might then select some scenario:

TRIMAG [t:\britz\t	rimag]			
File Settings	k			
-TRIMAG taks	-Scenario exploitation			
◯ Start TRIMAG	Scenario 1	my first scenario	*	
Exploit scenarios	Scenario 2	my test scenario	*	
	Promovio 2			
	acenario a			
	Scenario 4		×	
	Scenario 5		*	
		Show meta Show results		
TRIMAG Graphical User Ir	nter Ini file : gig.ini User name : unde	efined User type : runner		

And then, by pressing "show meta", view the settings used for these scenarios:

差 TRIMA	TRIMAG [t:\britz\trimag]					
View Handl	ng Windows					
🛃 TRIMA	📓 TRIMAG - result exploitation, meta data [0]					
				ो 🕨 🗘 🖸		
	my first scenario	my test scenario				
Scenario descriptior	my first scenario	my test scenario				
Choose model type	CGE	CGE				
Relative weight flows	30	30				
Use demand elasticities	false	true		ß		
Set substitutio elasticty	n 7.0	6.0				
Countries	NL,	NL,				
TRIMAG Gra	phical User Inter Ini file :	gig.ini	User name : undefined	User type : runner		

As the GDX container with the results comprises at the same time the numerical results and the meta-data, the information about the run is not lost as long as its result are not deleted. Sending the results to third parties or storing them on a SVN automatically also passes the meta-data along.

A specific format for META-data handling is available for CAPRI. It is shown in the following screen shot which also reports how the checkbox "Report SVN URLs and local modifications to include file" can be used in conjunction in META to make even more information available with regard to the run available:

🛃 Option					
Option					
User Settings CAPRI System	Settings GAMS and R SVN Other options				
SVN user id	•••••				
SVN password	•••••				
SVN URL for Gams	https://svn1.agp.uni-bonn.de/svn/capri/trunk/gams				
SVN URL for results	https://svn1.agp.uni-bonn.de/svn/capri/trunk/results/regcge				
SVN URL for restart					
SVN URL for data					
	$\overline{\mathbb{V}}$ Report SVN URLs and local modifications to include file				

As seen below, the include file will store the SVN URL for the different directories, the related revision number and will also report which files had local modifications:

too Anneloy not another
(set.countriesMeta).'Run scenario with market model'.'Run scenario with market model'.'svnRev modelDir' '3571'
(set.countriesMeta).'Run scenario with market model'.'Run scenario with market model'.'svnURL modelDir' '/svn/capri/trunk/gams'
(set.countriesMeta).'Run scenario with market model'.'Run scenario with market model'.'svnLocalMod1' 'T:\britz\capri_liaise\gams\supply\pmp_elas.gms'
(set.countriesMeta).'Run scenario with market model'.'Run scenario with market model'.'svnLocalMod2' 'T:\britz\capri_liaise\gams\arm\market_model_gms'
(set.countriesMeta).'Run scenario with market model'.'Run scenario with market model'.'svnLocalMod3' 'T:\britz\capri_liaise\gams\pol_input\liaise_liaise_scen.gms'
(set.countriesMeta).'Run scenario with market model'.'Run scenario with market model'.'svnLocalMod4' 'T:\britz\capri liaise\gams\capmod\set_and_store_dataout.gms'
(set.countriesMeta).'Run scenario with market model'.'Run scenario with market model'.'svnLocalMod5' 'T:\britz\capri_liaise\gams\capmod\oil_price_transmission.gms'

## File menu



# Settings menu

Settings		
	Edit settings	
	Load settings from ini file	
	Save current settings to ini file	
	Remove task specific settings	
	Remove view specific settings	

The settings dialogue was already discussed above.

# **Utilities and GUI menu**

Most of the utilities are discussed below in the following sections.

Utiliti	es	GUI	Settings Help	
	GDX Vie		ewer	
	C	SV to	GDX	
	St	tart C	ommand prompt	
	St	tart e	quation and variable viewer	
	В	uild H1	TML documentation	
	S	/N up	date	
GUI	Se	ttings	Help	
	Ba	atch e	xecution	
	Ge	enera	te GUI geometry from shapefile	
	Lo	ok an	d Feel	>

These options had been discussed above.

### Utilities: Batch execution

The batch execution facility is a tool which:

- Allows executing many different tasks after each other without requiring user input.
- Reports the settings used, any errors and GAMS result codes in a HTML page from which they may queried at a later time.
- Ensures that each new run generates its own listing file, which can be opened from the HTML page.
- Allows storing the output of the different runs in a separate directory, while reading input from unchanged result directories.

• Allows to use the GDXDIFF from GAMS to report automatically differences below a relative threshold of the generated results against preview ones.

The purpose of the batch execution facility is therefore at least twofold. On the one hand, it allows setting up test suits for the GAMS/R code of a project such as checking for compilation without errors for all tasks and different settings such as with and without market parts, etc.. Secondly, production runs of e.g. different scenarios can be started automatically. Timer facilities allow starting the batch execution at a pre-scheduled time. Along with functionalities to compare in a more or less automated way differences in results between versions, the batch facility is one important step towards quality control in GAMS/R based projects.

The batch execution allows starting a file, defining settings and tasks from the different work steps, e.g. form a modleling system such as CAPRI, and executing them without user intervention. Once started, the batch processor may be stopped so that the currently running GAMS program ends on its own ("end batch execution after next finalised GAMS step") or by sending a "CTRL-C" to the GAMS program. It will continue to run until the GAMS processor notices the CTRL-C – which may take a while – and then end with an error code. However, the GAMS processor will run some finalisation tasks as removing temporary files and directories.

#### Format of the batch execution steering file

- Generally, each line in the file comprises a keyword following by an equal sign and the related setting.
- Comment lines start with an asterisk. Blocks comment are between the keyword "ontext" and "offtext", which thus allows easily excluding blocks of lines from execution.
- The keyword "exit" prevents further processing.

#### Header

A batch execution file starts with a header which defines settings otherwise entered by user under settings dialogue of the GUI, i.e. directories, the GAMS engine the use and some further settings:

```
gams engine=d:\GAMS23.9\gams.exe
user = wolfgangb
work dir = t:\britz\capri\gams
 --- where the HTML page and the listings
      will be stored
output dir = d:\temp\batch
res dir = t:\britz\results
scr dir = d:\scrdir
gams options = scrdir=d:\scrdir --threads=on
 --- the following settings will write
×
           the results into different directories.
×
           The directory structure will be automatically
      generated
×
restartOutDir = t:\britz\capri\restart
resultsOutDir = t:\britz\results
number of processors = 30
processor speed relative = 160
number of iterations = 99
```

#### Settings for tasks

Each include file generated by GGIG comprises a block, commented out from use by GAMS, which can be pasted into a batch execution file, see example below:



The block of lines comprises also the minimum settings required for batch execution

(directories, GAMS executable).

Settings which do not change between tasks need not to be repeated, executing e.g. different scenarios is then simply done by changing the scenario file, followed by the keyword "execute" as shown below:

```
scen name=MTR_DECPL
execute=gamsrun
scen name=MTR_GREENGRAS
execute=gamsrun
scen name=MTR_GREENCROPD
execute=gamsrun
scen name=MTR_GREENsET
execute=gamsrun
scen name=MTR_GREEN
execute=gamsrun
scen name=MTR_CONV
execute=gamsrun
```

#### Notes:

- *execute* = *gamsCompile* will never execute the program, if this option is switched on the interface (see below). This is useful to define tests which comprise some cases to run (optionally) and other for compilation, only.
- *execute* = *gamsTestCompile* will run automated test for most control. These tests build on the default setting for all controls on the interface. In a loop, for each single selection control such as a checkbox, all potential settings are subject to a test, documented on the HTML page. Afterwards, the control is reset to its default and the settings for the next control are tested. Controls which define numerical input (such as sliders or spinners) are normally not subject to such tests, as introducing a different number in the code is unlikely to provoke errors in compile time tests. It is however possible to define, besides the default registered with the control, a second numerical value for a test (*test*= ... as a style attribute of the control). This is useful if the code treats, for instance, a zero different from a non-zero value. Equally, controls can be excluded from testing (*test=false* as a style attribute of the control.
- The use the *task* keyword will let the batch routine load default values for all controls, overwriting any setting inputted before.
- This also means that it is not necessary to specify inputs for all settings. Only non-default values need to be inputted. To help here, the section with the task

specific settings in the generated include file is split in two. First, the non-

defaults are listed:

```
task = Single farm run
* -----
  Non-Default settings in use
                       _____
 Scenario description = enterName2
 Farm branches = Arable, Biogas
 Post-Model stochastics = true
 CropsSel = WinterWheat, WinterBarley, WinterRape, Summerpeas, Summerbeans, Idle
 Organic farming = enforced
 Allow manure import = true
 Restriction from German Fertilizer directive = false
 Stop point = After first RMIP solve
 Use RMIP = true
 Use heuristics to reduce number of binaries = false
 Use heuristic fixing of binaries = 0
 Additional reporting = true
```

And next a list with all default, with the non-defaults being commented out:

```
All available settings, settings at their defaults as comments
* --
 Scenario description = enterName2
 Farm branches = Arable, Biogas
  Sets = sets de
  Machinery = mach_de
* Feeds = feeds de
  Silos = silos de
 Buildings = buildings de
 Stochastic programming = false
 Post-Model stochastics = true
 Allow for off farm work = false
  Allow hiring full-time farm workers = false
  Allow buying or selling land = false
  Max land bought per plot [ha] = 10.0
  Allow leasing land = false
  Land rent [Euro/ha] = 310
  Calibrate = false
  File with calibration results = farm empty
  Dynamics\n in model = Comparative-static
```

- If switching between tasks is required and control settings should not get lost, use *taskLoad* instead.
- Controls settings might be updated based on dependencies.

#### Program flow possibilities in the batch execution facility

The user can, similar to GAMS, define string macro:

MACRO variant = MTR

Which can be used in following statement for string substitution, e.g.

Scenario = %variant%

Settings for controls are automatically stored as macros.

The batch execution facility offers the following simple scripting possibilities:

For – endfor

The "for" loop allows to automatically set a range of string values to a macro, separated by a forward slash, e.g.

```
For variant = MTR_SIMPLE / MTR_ADVANCED
Scenario = %variant%
BatchName = %variant%
execute = gamsrun
Endfor
```

Nested for loops are allowed.

*If*-elseif-else-end

If blocks allow executing parts of the scripts conditioned on string comparison, a feature is basically only necessary in conjunction with loops. Please note that nesting of if-statements is not allowed. However, similar to the "for" statement, and if statement can match true to several right hand side strings (if  $\% x\% = y1 \setminus y2 \setminus y3$ ).

```
if %variant% = Fixed Allocation of global investments
   Foreign Savings = Fixed global allocation of investments
   hide_Modules = CO2_Emissions
elseif %variant% = Global Bank
   Foreign Savings = Global equal returns to capital
   hide_Modules = CO2_Emissions
else
   Foreign Savings = Global equal returns to capital
   hide_Modules = GTAP_AGR,GTAP_AEZ,CO2_Emissions
endif
```

#### Parallel execution

The user can start GAMS tasks between the two keywords

startParallel

CollectParallel

which will be executed in parallel. Further execution will wait at the "collectParallel" point until all parallel tasks have finished. The code will ensure that GAMS tasks are only compiled once the previous task has been started, in order to avoid problems with file processing at compile time. Equally, sub-directory in the "scratch" directory will be generated and passed to the GAMS processes to avoid parallel I/O. However, the user must make sure that the GAMS code of the parallel tasks does not access the same files in write access as that might lead to unpredictable program behaviour.

#### Using the batch execution facility

The batch execution utility can now be opened from the menu bar under "GUI":

1		
GU	I Settings	
-	Batch execution	F
3	Generate GUI geometry from shapefile	F
onte		

It will open as a separate window as shown below.

#### Graph: Batch execution panel



If the suite of tasks comprises "execute" statements, those can be downgraded to "compile" with "Only compile GAMS programs" check box.

The check box "Generate EXP/REF files for HTML documentation" adds settings to the GAMS calls which generate two specific reference files by the GAMS compiler which comprise information of files and symbols used by GAMS. For details on the code documentation facility see the technical document "Javadoc like technical documentation for CAPRI" to be found on the Capri web page under technical documents. The "directory for exp/ref files" defines where those files will be stored.

The batch language allows definition of a timer, i.e. to start the execution at a specified time.

#### The output from batch execution

As it is assumed that batch execution will not be monitored by the user during execution, a logging mechanism is established. Listing files and generated include files are stored in subdirectories of the "output dir" defined in the batch execution file:

```
* --- where the HTML page and the listings
* will be stored
*
output dir = d:\temp\batch
```

The sub-directory is named after the time point (date followed by time including secons) where the batch execution is started:

4.07.2014_07.34.12	24.07.2014 08:57	Dateiordner
4.07.2014_07.40.32	24.07.2014 08:57	Dateiordner

These sub-directories comprise the listing files and generated include files, as well, if requested GDX files with result differences from the Gams GDXDIFF utility, labelled according their starting sequences, e.g.:

▲ 0_1.lst	24.07.2014 07:53	LST-Datei	158 317 KB
▲ 1_1.lst	24.07.2014 08:06	LST-Datei	158 403 KB
▲ 2_1.lst	24.07.2014 08:19	LST-Datei	172 658 KB
▲ 3_1.lst	24.07.2014 08:32	LST-Datei	172 673 KB
▲ 4_1.lst	24.07.2014 08:45	LST-Datei	156 906 KB
▲ 5_1.lst	24.07.2014 08:57	LST-Datei	955 KB
🞒 batch.html	24.07.2014 07:53	HTML-Dokument	15 KB
fortran.gms.0_1	24.07.2014 07:40	0_1-Datei	18 KB
fortran.gms.1_1	24.07.2014 07:53	1_1-Datei	18 KB
fortran.gms.2_1	24.07.2014 08:06	2_1-Datei	18 KB
fortran.gms.3_1	24.07.2014 08:19	3_1-Datei	18 KB
fortran.gms.4_1	24.07.2014 08:32	4_1-Datei	18 KB
fortran.gms.5_1	24.07.2014 08:45	5_1-Datei	18 KB

Additionally, a HTML page reports all tasks which have been started, the return code of the GAMS process and all major setting, as well as link to open the listing file with the editor. The following screen shot shows the first part of the HTML page resulting from executing a batch file. Tasks which did yield a non-zero GAMS return code and errors are shown in red.

GGIGRI batch execution report										
Runs are generated from					d:\temp\12.02.2015_07.38.57\batch.html					
at					Do, 12 Feb 15 07:38:57					
Workstep	Task	User	Directories: work result dat restart	Settings	etings		Started Ended Time used	RC	Listing GdxDiff	
Single farm runs	Single farm run		U:\dairydyn\gams \results null null	Dairy, test with 60 cows Updated settings( <u>click</u> All settings in use( <u>click</u>	and 2.5 Ak, until 2020, no arable crops to anountroled to anountroled	run ⊷scen≋incgen\runInc ⊸ggig=on	Do, 12 Feb 15 07:38:57 Do, 12 Feb 15 07:39:43 0 min and 45 secs	0	D:temp112.02.2015_07.38.5710_1_list D:temp112.02.2015_07.38.571dff_0_gdx p_sumRes # of diffs > 1.0E-10% : 8	
Single farm runs	Single farm run		U:\dairydyn\gams \results null null	Dairy, test with 60 cows Updated settings( <u>click</u> scenariodescription = 60 farmbranches = Dairy,A All settings in use( <u>click</u>	airy, test with 60 cons and 2.5 Ak, with anable crops gladed settings( <u>cick</u> to <u>absuit)field</u> metanolescoptor - for <u>Cons</u> _ SAAs, Dairy_Anable metanolesc = Clany, Acade, settings in well( <u>cick</u> to <u>showhide</u> )		Do, 12 Feb 15 07:39:43 Do, 12 Feb 15 07:42:58 3 min and 15 secs	0	D:temp112.02.2015_07.38.57/1_1 lst D:temp112.02.2015_07.38.57/diff_1_gdx p_sumRes # of diffs > 1.0E-10% : 16	
Single farm runs	Single farm run		U:\dairydyn\gams \results null null	Arable, test with 2.5 Ak, Updated settings(click scenariodescription = 2 farmbranches = Arable, crops = WinterCere, Sun All settings in use(click	all anale crops Lo bonchide SAs_Anable mera/Cee, Warden Rope Potatoes SugarBeet MacCom, SummerPees SummerBeans, Ide, MacCOM, MacSil, WheatGPS, Lo abonchide)	run scen≍incgen\run1nc ggig=on	Do, 12 Feb 15 07:42:58 Do, 12 Feb 15 07:43:09 0 min and 10 secs	0	D:templ12.02.2015_07.38.57\2_1_lst D:templ12.02.2015_07.38.57\dff_2_gdx p_sumRes # of diffs > 1.0E-10% : 12	
Single farm runs	Single farm run		U:\dairydyn\gams \results null null	Dairy, test with 60 cows Updated settings(click farmbranches = Dairy, A crops = WinterCere, Sun restrictionfromgermanfe scenariodescription = 60 All settings in use(click	and 2.2 As, with installer cripts, no fertilizer directive to broathdidu abble. Directoscolor = for the second second second second second second mechanisms and second second second second second second second Directoscolor = for the second second Directoscolor = for the second sec	run scen≈incgenirunInc ggig=on	Do, 12 Feb 15 07:43:09 Do, 12 Feb 15 07:45:56 2 min and 47 secs	0	D.templ12.02.2015.07.38.57\3.1.lat D.templ12.02.2015.07.38.57\diff.3.gdx p_sumRes # of diffs > 1.0E-10% : 2	

The table cell with the task settings can become quite long:
#### **GGIGRI batch execution report**

Runs are genera	ated from				d:\temp\12.02.2015_07.38.57\batch.html	
at					Do, 12 Feb 15 07:38:57	
Workstep	Task	User	Directories: work result dat restart	Settings		Gams optic
				Darry tast with 60 cows Updated settings in ussci(icit) Scenario description '8) 'Farm branches' Dairy'. 'Different states of natur 'Allow for off farm work'. Allow leasing land 'fals 'Allow buying land 'tals 'Allow buying land 'tals 'Allow buying land 'tals 'Allow buying land 'tals 'Last Year' 2010' 'Last Year' 2020' 'Dynamics in model' 'Ful 'Ime resolution for fined 'Ime resolution for fined 'Construction year of sta Household consumption 'Growth rate of househou 'Growth rate of househou 'Conv breeds' HF; 'Milk yield (in 100 kg/Co Genetic dynames of stalg Allow selling of raisage 'Allow selling of raisage 'Allow selling of raisage' 'Allow selling of heifers' 'Crops' WinterCere Sum 'Allow selling of heifers' 'Crops' WinterCere Sum 'Nomiton' falsage' talse' 'No tillage 'talse' talse' 'No tillage 'talse' talse' 'No tillage 'talse' 'Last'	and 2.5 Ak, until 2020, no arable crops to abcombined DCows _2_5Aks_Dairy' 6 of price "false" false " y' (default]' hous investment (Euro /ha or hour)' 3.0 [default]' (by anamic without prolongation' strendtoff farm labour decisions' 1 (default]' use '3 (default]' bles resp. age of machinery '1995 [default]' if consumption in % p.a.'1.0' if consumption in % p.a.'1.0' '0 [default]' 0' w)' 30' i' 0 [default]' 1' 0 [default]' 1' 0 [default]' 1'' 1' 0 [default]' 1'' 1'' 0 [default]' 1''' 1''' 1'''''''''''''''''''''''	

such that some simple JavaScript allow to close / hide these settings.

```
Norway Dairy
Updated settings(click to show/hide)
All settings in use(click to show/hide)
Non default settings in use(click to
show/hide)
```

There are three blocks with such settings:

- *1. Update settings:* These are the control settings which had been changed compared to the previous batch run.
- 2. *All settings in use*, as the name suggests, these covers all control settings currently active for the task. Some might be taken out; depending on how GGIG is used this can refer to controls currently invisible.
- *3. Non-default settings in use:* same list as under 2., but without settings equal to default ones.

The hide / unhide of the settings is based on the following:

hideSettings = none,use

The standard setting will hide all settings when the page is opened. "batch" refers to the settings explicitly entered before a task is executed, "use" to all settings referring to the task. The statement above thus first would show all settings, and next hide those in use, but not those explicitly entered in the batch.

#### Automated GDXDIFF

#### The GDXDIFF utility from GAMS

(http://www.gams.com/help/infocenter/index.jsp?topic=%2Fgams.doc.tool%2Fgamside%2Fu tilgdxdiff.htm) allows comparing one or several GAMS symbols in two GDX files. It can be used with batch utility to compare the results from the currently run task against pre-existing ones. The utility generates a batch file which reports all entries beyond a pre-defined relative or absolute difference, in here, only relative ones are implemented. In order to use the utility, the user must indicate which GDX file and symbol should be compared:

gdxdiff=gdxfile=res\_60Cows\_2\_5Aks\_Dairy\_Arable\_until\_2020,id=p\_sumRes,relEps=1.E-12

The three keywords are hence:

- 1. **gdxfile**: the name of GDX file where the results to be compared are found. The result directory fom the task is used to locate the file.
- 2. id: the GAMS symbol to compare
- 3. **relEps**: the minimum relative difference used for reporting

The GDX entry in the HTML page lists the name of the generated file (it can be opened with a GDX file viewer) and the # of entries found above the threshold:

c.\scrain(17.01.2021_10.18.00\alii_2.gax p_sumRes # of diffs > 1.0% : 23(click to show/hide)
WinterWheat dif1 40 60
WinterWheat dif2 36 46
WinterRape new 16 40
profit dif1 29278 44
profit dif2 208230 15
NSurplus dif1 57 50
NSurplus dif2 10 17
DSurplus diff 12.69
PSurplus dif2 0.22
rouplus, dif1 1091 50
totalLabour, dif1, 1081, 50
totalLabour, dif2, 7200,00
labourUse,dif1,0,15
labourUse,dif2,1,00
margArab, dif1, 1123, 84
margArab, dif2, 962, 46
MaizSil,old,20,54
gras2_12_graz100,old,4,13
gras4_14_4cuts_sil100,old,35,87
cows,old,138,51
gras,old,40,00
NorgAppl,old, 162,49
labourBoundMonthly.old.3.00
margGras.old.740.29
en en anti Marca en esta sut de la constante en anti-

As seen above, similar to the settings, these differences can be inspected directly in the

HTML page.

## Error reporting and outputting results from GAMS to the HTML page

Batch run which results in a non-zero return code from GAMS are flagged in red:

		Single farm runs	Single farm run		C:\farmdyn\gams \results \dat T:\britz\dairydyn \restart	Post-Model stochastics=true Updated settings(click to show/hide) All settings in use(click to show/hide) Non default settings in use(click to show/hide)	compile scen=incgen/runInc ggig=on procDirPath=c:\scrdir optdir=opt	Fr, 19 Mär 21 14:45:32 Fr, 19 Mär 21 14:45:34 0 min and 2 secs	2 Errors (click to show/hide)	<u>C:\temp</u> \19.03.2021_14.45.00 \11.lst
--	--	---------------------	--------------------	--	--	---	---	---	--	---

The errors which occurred can be directly inspected by the "click to show/hide" in the return code column:

L		1	1	1		1	1
Single farm runs	Single farm run	C:\farmdyn\gams \results \dat T:\britz\dairydyn \restart	Post-Model stochastics=true Updated settings( <u>click to show/hide</u> ) All settings in use( <u>click to show/hide</u> ) Non default settings in use( <u>click to show/hide</u> )	compile scen=incgen/runInc -ggig=on procDirPath=c:\scrdir optdir=opt	Fr, 19 Mar 21 14:45:32 Fr, 19 Mar 21 14:45:34 0 min and 2 secs	2 Errors (click to show/hide) *** Error 349 in C:Varmdyn UgamsVæptolier/biodiv_ind_salca.gms Declaration not allowed inside a LOOP or IF statement *** Error 194 in C:Varmdyn UgamsVæptolier/biodiv_ind_salca.gms Symbol redefined - a second data statement for the same symbol *** Error 257 in C:Varmdyn UgamsVæptolier/boat_model_stoch.gms Solve statement not checked because of previous errors	<u>C:\temp</u> \19.03.2021_14.45.00 \111.lst

Equally, clicking on the listing will open it in the browser to gain further insights.

```
(\leftarrow) \rightarrow C \land
                  file:///C:/temp/19.03.2021_14.45.00/11.lst
                                                              F
                                                                  ..
GAMS 34.3.0 rac355f3 Released Feb 25, 2021 WEX-WEI x86 64bit/MS Windows - 03/19/21 14:4
General Algebraic Modeling
                                         System
Compilation
  1 *****
       FARMDYN project
      GAMS file : EXPSTARTER.GMS
       Opurpose : Top level program of single farm runs
              : W.Britz, B.Lengers
       @author
      @date
              : 12.11.10
       @since
              : 2010
       @refDoc
              :
       @seeAlso
              :
       @calledBy :
 18
    *
 19
38756
      set dummyGras / gras4_14_4cuts_sil100,gras2_12_graz100 /;
****
                $194
       $349
**** LINE
                      C:\farmdyn\gams\exploiter/biodiv_ind_salca.gms
         500 BATINCLUDE
                      %2 "draw"
```

It is also possible to pass directly output from GAMS to the HTML page generated by the batch utility as seen in the following example:

Violatzuanyuyii vionalait setungs in userunta to Viestart show/hide) optim=opt 0 in constant -c.isudui 0 min and 3 secs MaizSi : 0.01 Comustard : 0.01 Fattres : 1125.00		Single farm runs	Single farm run		C:\farmdyn\gams \results \dat T:\britz\dairydyn \restart	Fattners, 10 ha land, manure export Updated settings(click to show/hide) All settings in use(click to show/hide) Non default settings in use(click to show/hide)	run scen=incgen/runInc ggig=on procDirPath=c:\scrdir optdir=opt	Fr, 19 Mär 21 13:00:32 Fr, 19 Mär 21 13:00:35 0 min and 3 secs	0	C:\qmTemp\19.03.2021 12.59.5 \d.lst profit : 75561 WinterWheat : 6.67 WinterBarley : 1.33 WinterRape : 1.99 MaizSi : 0.01 CCmustard : 0.01 Fattners : 1125.00
---	--	---------------------	--------------------	--	--	--	--	--	---	---

This is useful if batch files are used for repeated tests. An experienced model use will be often able to conclude from key results if there had been major problems with the test. The following GAMS code shows how this can be done. Any line found in the log which starts with 'html' will be stored by the batch routine and added to the HTML page. In GAMS, put\_utility 'msglog' allows to write to the log file, and this can be used to output selected results as documented below.

```
*
*
* --- write the profit, acreages and herd size to log, and to HTML page if the natch mode is running
*
put_utility batch 'msglog' / 'html profit : ',p_sumres("profit"):0:0;
loop (curCrops $ (p_sumRes(curCrops) gt 0),
    put_utility batch 'msglog' / 'html ',curCrops.tl:0:0,' : ',p_sumres(curCrops):0:2;
);
$$iftheni.possHerds defined possHerds
loop (possHerds $ (p_sumRes(possHerds) gt 0),
    put_utility batch 'msglog' / 'html ',possherds.tl:0:0,' : ',p_sumres(possHerds):0:2;
);
$$endif.possherds
```

## Utilities: Generate GAMS documentation in HTML pages

#### -GAMS documentation generation Directory with input files t:\britz\capri\codeDocInput Set directory Set directory Directory for HTML documentation files t:\britz\capri\GAMSDoc t:\britz\capri\codeDocInput\Baseline\_calibration.ref t:\britz\capri\codeDocInput\Build\_global\_database.rel t:\britz\capri\codeDocInput\Build\_HSMU\_database.ref t:\britz\capri\codeDocInput\Build\_regional\_database.ref t:\britz\capri\codeDocInput\Finish\_national\_database.ref List of available EXP and REF files t:\britz\capri\codeDocInput\Generate\_expost\_results.ref t:\britz\capri\codeDocInput\Generate\_policy\_shifts.ref t:\britz\capri\codeDocInput\Generate\_trend\_projection.ref t:\britz\capri\codeDocInput\Prepare\_national\_database.ref t:\britz\capri\codeDocInput\Run\_simulation.ref < > Generate HTML documentation

Graph: Panel to steer GAMS documentation generation

The GUI comprises a tool to generate for each GAMS file and each symbol used HTML pages which are interlinked. For details on the code documentation facility see the technical document "Javadoc like technical documentation for CAPRI" to be found on the Capri web page under technical documents.

The controls on top allow the user:

- To define in which directory the "EXP", "REF" and "GDX" files are stored which serve as input into the documentation generator.
- To choose the directory where the HTML files will be generated.
- To select the tasks covered by the documentation generator.

#### Structure of the HTML pages

There are basically two types of HTML pages:

- 1. *Pages for individual objects* (parameters, sets, variables, equations, models, acronyms, functions, files and source files)
- 2. *Summary pages for classes of objects*, per project in alphabetical order. An additional page lists all set elements.

The pages for the individual objects carry the following information:

- Name of the object (e.g. DATA) and type (parameter, set, variable etc.)
- Long text description as given in GAMS declaration
- Domain information, as hyperlinks to the domain sets.
- In which files and for which projects (as capmod, capreg ...) the object is declared, defined, assigned and referenced.
- In the case of sets: derived subsets, and objects where the set is used as a domain. Elements of the sets and the subsets.
- In the case of source files: which symbols are declared, defined, assigned and references in the files. Information from SVN (version, local modification,s out-of-date with server). Included files, and files which include the file. For GDX files: where included and included by which file.
- "Tagged" in-line comments taken from the source code files, what is called "doclet" (see e.g. <u>Sun document about how to write Doc comment for</u> <u>JavaDoc</u>) in JAVADOC, see .e.g. <u>wikipedia article</u>

#### **Tagged in-line comments**

Similar to the element comments underlying JAVADOC (see e.g. ), "tagged" in-line comments are proposed for the inline code of CAPRI (sometimes called "doclets", e.g. ). The following shows a possible implementation which is currently already operational:

•••

```
* @start
```

```
* @author W.Britz
```

- \* @docRef perfect aggregation of production
- \* @seeAlso gams\capreg\cons\_levels.gms

```
MODEL CONS_LEVLS / ...
```

```
••
```

In the example above, the REF file will comprise the information were the model CONS\_LEVL will be declared, and the JAVA application will search backwards for lines

with tags (@..). Those tags will be linked to the object, and integrated in the HTML pages. The @start tag must be used to declare the start of the documentation for the current symbol.

#### **Refactoring Consequences for Gams Code**

- All files should carry a header which reports the purpose of the file, and if possible, an author (contact person). The file header should start with a line of stars and end with a line of stares. All lines in the file header should start with a "\*".
- 2. The use of *\$GDXIN* is discouraged as it may load in huge amounts of data at run-time. Equally, it will load element codes comprised in the data sets even if they are not referenced later in the code. The only exemption is when the symbol must be loaded at run time as in case of META data, instead, execute\_load should be used.
- 3. An "\$IF NOT EXIST *myFile* \$ ABORT *myFile* is missing" statement should be in the line before "execute\_load *myFile someSymbols*".
- 4. An "\$IF EXIST *myFile* \$ LOG *myFile* will be overwritten" statement should be in the line before execute\_unload *myFile someSymbols*".
- 5. All symbols should be declared with a clear long text description, i.e. statement in the style "SET A;" are discouraged.
- 6. Code in lengthy files should be moved into new files which are included so that a more modular structure is evolving. The new file should have a clearly defined and encapsulated task which is described in the file header.
- 7. Symbol declaration should where necessary be preceded by a "doclet" of the form
  \* @start or, alternative, a blank line
  - \* @DocRef reference to the methodological documentation (optional)
  - \*@ seeAlso reference to other file or symbol (optional)
  - \* Any comments

Declaration (as SET A "The alternative technologies per production activty" / T1,T2 /;

8. Symbols, especially when they are not widely used across programs should carry meaningful names.

Other recommendations arising from analysing the files are:

1. Single lines in the code should not exceed the size of a normal screen width when using medium sized fonts.

- 2. Indentation should be used to render the program structure defined by loops, if statements and the like more visible.
- 3. Especially tricky statements which use complex \$ operators, several cross-sets and the like should be preceded by some explanatory comments.
- 4. Symbols which are only used locally in a file should be deleted from memory by "option kill= …".<sup>4</sup>.
- 5. Before defining a new set one should check if not the very same collection of elements is not already defined.
- 6. Lengthy data tables should be moved into a gdx file to reduce the number of code lines.
- 7. Data should be accompanied by meta data.



<sup>&</sup>lt;sup>4</sup> A feature request was sent to GAMS to support local scope, so that a symbol can be declared local for a file and subdirectory, and the compiler will raise an error when it is used out of scope.

## Example for a Symbol page



## Example for a GamsSourceFile page



## Example for a page for the a set



#### Set element list



## Utilities: Equation and variable viewer

## **Background and motivation**

Complex GAMS code and related models such as the market model of CAPRI with its ~70.000 equations and variables are very hard to debug. The model listing produced by GAMS from such a model is quite long, and filtering out e.g. all lines belonging to a certain market not possible with all editors. Further on, linking the listing to the equation structure of the model is also far from easy. That paper describes a utility linked into the GGIG, the GUI generator used by CAPRI, which supports working with large model outputs (and more generally complex GAMS projects with many symbols).

The new tool also incorporates the functionality of the existing GDXViewer comprised in the GUI. The tool can be used to "track" changes to symbols in the GAMS code by producing a range of GDX files (currently up to 5) at different execution points.

#### An overview on the viewer

S GDXViewer												- • ×
File Views										_		
A Selection										T a Y	2 Options	
Selection											Ontions	
Section				1.1-	1.0	1-0					(options	
	Variables *	✓ Equal	stons *	▼]1]*	▼] <sup>2</sup> [*	▲ 3.	• •				Set XML table definition	file
E Lice indents in equation ou	they at	v_GLDemandFS(AFR_LDC)	4		ArmBal1_(AF	FR_LDC, AP	PL)	<u> </u>			Use table definitions from : nu	
	Variable selection	v_GLDemandFS(AFR_REST)		Equation	selection ArmBal1_(AP	FR_LDC,BA	RL) ==)				📰 Sort code lists	
Use small font for non-sele	ected var terms	v_GLDemandFS(ALG)	-		ArmBal1_(AF	FR_LDC,BIO	D)	-			📄 Show dialog to link GDX dimen	sions to sets
	<i>.</i>					~					<u> </u>	
🖆 Gams includes 🛛 📼	🛃 Gams					📕 🛃 Syn	nbols in GDX					
🕌 T:\britz\capri\gams\capmod.gms 🔄	T:\britz\capri\gams\capmod.gm	s \gams \sets.gms				List	of tables load	led from GDX file(s)				
\gams\util\acronyms.gms								T:\brit	z\capri\gams\convert	market	_model_2.gdx	
	Contrast						user input	name	type	dims	records long text	
\gams\util\title 1.gms	VOICEAC						<u> </u>	lines	set		1 10001	*
• • • \gams \util\kil_model.gms	CAPPT project							dummyTrd	set		1 1	
<ul> <li>(gams)capmod/define_regional_sets.(E</li> <li>)asmolarm/arm sets arms</li> </ul>	Carine project							META_STEP	set	-	1 1	
	GAMS file : CADM	IOD. GMS					<u> </u>	META	set	-	4 16401	
\gams\regcge\regcge_decl.gms	SANS LILL . CAPE	was a second						SimulationYear	set		1 1	
\gams \regcge \regcge_capri.gms	Anurrosa I Ton	level of the CAPPT air	wlation engine					Countries	set	_	1 35	
\gams\reports\rep_sets.gms	epurpose . rop	Tever of the Obert Sta	diation engine					ntsivi	parameter		0 1	
ganstervirdigets_anno.gns    oansteed\feed ded.ans	Southor IN R	with a Goobt M adv	nanner T. Janagon et	1				SCEN	set	-	1 2	
\gams\supply\def_supply_model_par.	Adata 1 22 1	0 00	indeuer, r. oansson et				4	VEADO	ant		al al	
+\gams\capmod\create_sim_ini_gdx.gn	Batron 1999	0.03										
\gams\capmod\load_sim_ini_gdx.gms    ams/cupplu/usid_alas.gms	BrefDoc :							T:\brit	z\capri\gams\convert	market	model 1.adx	
	Reeableo :						user input	name	type	dime	records long text	
\gams\inputs\yieldIY.gms	(DCCM2DD 1						user input	lines	pet	unis	1 10001	_
•\gams\capmod\define_inputs.gms	Sofftaxt							dummyTrd	set		1 1	
\gams\regcge\regcge_templ.gms    amslregcge\lead.unlead.ini.gms	**********							META_STEP	set		1 1	
<ul> <li></li></ul>	Soffligting							countriesMeta	set	-	1 35	
•\gams\sugar\breakdown_quots.gms	Sofflog							SimulationYear	set	-	1 1	
\gams\capmod\inflation_and_trend_in	SSTARS SSSS							Countries	set		1 35	
	Saetglobal metawh o	ff						ntsM	parameter		0 1	
(gams/policy/def policy.gms								SCEN .	set	-	1 99	
\gams\policy\del_policy.gms     +	* save/restart	point for debugging pu	proces in the market	model calibra	tion -		-	VEADO			1 1	-
	1	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		Y	-	-						
Leguations					🛃 Variables							
ArmBall (AFR LDC, APPL) "Adding	up of human consumpt.	ion feed and processin	g "		v consQuant (A	FR LDC,	APPL) Human	consumption [1000 tons	1.1=87,126.1o=0	upp=8	7.125,678	
					-	-						
-0.0114645139135529*v_consQue	ant (AFR_LDC, APPL)											
+0.0114645139135529*v_arm1Qu	ant (AFR_LDC, APPL) =E=0	9										

The viewer comprises a number of windows

- Selection: A window where variables and equations can be selected, and filters for variables and equations defined
- **Options**: A window with options for the GDXViewer
- Gams includes: A tree view of the GAMS includes
- **GAMS**: A window showing the GAMS code
- Symbol from GDX: A view on a current selected GAMS symbol
- Equations: A view on the current selected equation(s), in linearized form
- Variables: A view on the currently selected variables

The windows can be dragged, resized and minimized.

#### Producing input for the view with GAMS

The viewer can be used in different configurations, which are available via the "File\run" menu:

File Views	
Exit	
Run 🕨	Load only GDX files into viewer
	Load GAMS files (and if existing GDX files) into viewer
	Load convert output into viewer
	Load convert output and GAMS file into viewer

• "Load only GDX files into viewer" comprises basically the same functionality as the existing GDXViewer. However, the selection panel and the view on the symbol are both visible at the same time.

🔬 GDXViewer										
File Views							_			
				_			0	( - Y		
Selection								- 0	Options	0
Selection									Options	
	SSQ.	<u>^</u>	ArmBa	1_	<u>^</u>				Set XML table defin	nition file
	Variable selection	Eq	uation selection	1_					Use table definitions from	n : null
	Use small font for non-selected var terms vo ov bevFunctor	taddFac	RioPro	4 Maroáon					Sort code lists	
									Show dialog to link GDX d	imensions to sets
Sumbol from GDY			Symbols in GD	(						
Countries			List of tables k	aded from GD	() file(s)					
- countries				ouco nom op.	A metal					
		ो <b>- २ र</b>			T:\britz\ca	apri\gams\convert_	_market_	model_2.g	dx	
				user input	name	type	dims	records	long text	
					lines	set	1	1 1000	1	6
¥					META STEP	set		1	1	
BL000000		Belgium and Luxembourg			countriesMeta	set		1 3	5	
DK000000		Denmark			META	set		4 1640	1	
DE000000		Germany			SimulationYear	set		1	1	
EL000000		Greece			Countries	set		1 3	5	
E \$000000		Spain			risivi	parameter	-	1 0	4	
FR000000		France			SCEN	set		1 ,	2	
IR000000		Irland		41	MEANE	+			4	*
1T000000		taly								-
NL000000		The Netherlands			To both a				a	
AT000000		Austria			1: (Dritz (Ca	ipri (gains (convert_	market	model_1.g	ux	
PT000000		Portugal		user input	name	type	dims	records	long text	
SE000000		Sweden			dummyTrd	set		1 1000	1	6
F1000000		Finland			META STEP	set		1	1	
UK000000		United Kingdom			countriesMeta	set		1 3	5	
CZ000000		Czech Republic			META	set		4 1640	1	
HU000000		Hungary			SimulationYear	set		1	1	
PL000000		Poland			ntsivi	parameter		2	1	
\$1000000		Slovenia			step	set		1 9	9	
SK000000		Slovak Republic			SCEN	set	1	1	2	
EE000000		Estonia			MEADE	land			AI F	
LT000000		Lithuania								
LV000000		Latvia								
CY000000		Cyprus								
MT000000		Mata								
BG000000		Bulgaria								
R0000000		Romania								
NC000000		norway								
TUR		Abasis								
AL000000		Albania								
MK000000		Macedonia Cashia								
C 5000000		Scibia								
MO000000		Constin								
HR000000		Croata								
BA000000		posnia anu rierzegovina								
K0000000		NUSUYU								
		< >								

 "Load GAMS files (and if existing GDX files) into viewer": With a GAMS file and one or several GDX file: browse the source code and click on highlighted symbols to load them from the GDX file(s). That is basically a GDX viewer linked to the GAMS code. The GDX files can be produced in GAMS with "execute\_unload someName.gdx;" which will dump all GAMS symbols into a GDX file.

GDXViewer	The second se	-	No. of Lot of Lo	
File Views				
Selection			- 0	🕼 Options 📃 😐
Selection				Options
🔄 Use	indents in equation output VB Arm	nBal1_ nFit1	\$	Set XML table definition file
Use	small font for non-selected var terms VD Equation selection Arm	nFit2_		Use table definitions from : null
	pv_bevFuncIntAddFac + BioF	ProcMarg/	Agg_ +	Sort code lists
				Show dialog to link GDX dimensions to sets
📓 Gams includes 💿 🖷	🕼 Gams 🖂		Symbol from GDX	
🕌 T:\britz\capri\gams\capmod.gms	T:\britz\capri\gams\capmod.gms t:\britz\capri\gams\util\acronyms.gms		NCNC	
	Sinclude 'capmod\define regional sets.gms'	^		ो 🕂 🕹 🐱
			: •	adv
	<pre>set debugRegions(Rall);</pre>			jux
	option kill=debugRegions;		N CAL Nut	rient content, calories
	* load gate yead by market model		N_FAT Nut	rient content, fat
	* Toda sets ased by market moder		N_PRO Nut	rient content, protein
	<pre>\$include 'arm\arm_sets.gms'</pre>			
	*			
	* policy set definitions and regional small producer shares			
	Sinclude 'policy\policy sets.qms'	=		
	Siftheni %regcge% == on			
	* the declaration part of regional CGPs			
	*			
	Sinclude 'regoge_decl.gms'			
	*			
	p_firstCallCGE = 1;			
	* declaration of links to CAPRI			
	· ·			
	<pre>\$include 'regcge\regcge_capri.gms'</pre>			
	* (maMSIng)(comMSWork) S = Ing)(comMSWork) = VPS.			
	*			
	Şendif			
	*			
	* set definitions for different type of reports			
	SET RSS (RALL);			
	SET RSST(RALL);			
	Set NCNC(COLS) "Common nutrient coefficients" /SET.NCNC_POS / ;			
	Sinclude 'reports\rep_sets.gms'			
	ALIAS (STMY, STMY1):			
	SET SIMYY1;			
	· ·			
	ALIAS (STEP, STEP1, STEP2);	Ŧ		

• "Load convert output into viewer": with Convert output only. Convert is a "solver" shipped with GAMS which generates a linearized version of a model with obfuscated variable and equation names, e.g. to ship it to a solver developer for testing. Convert can produce a dictionary file which allows to link it to the obfuscated names.

DXViewer							
Views							
ielection					o ís	Options	
election						Options	
Variables <sup>®</sup> Provations <sup>®</sup> Variables	- Equations *	_] 1(*]			L C	Sat YML table definiti	ion file
Vanaura Luquara Vanaura	• Equilation	• • •			L L	Set AFIL table definition	Johne
pv_bevPuncSubsExpCorrPact	ArmBal1_		<b>^</b>			Use table definitions from :	null
Variable selection	quation selection					Sort code lists	
Use small ront for hon-selected var terms pv_selectors pv_feedConv +	BioProcMargAg	29_	-			📰 Show dialog to link GDX dim	nension
Aumbal from GDY	C L Sumbols in GDV				_		
rm2Val_	List of tables loaded	I from GDX file(s)					
Dim2 (Lev/,Marginal,)		T:\	britz\capri\gams\convert n	narket mod	el 2.gdx		
Levi Table		er innut name	type	dime reo	vede la	ona text	
		ines	set	1	10001	A	
BIOE BIOD RTEM BARL DATS MAIZ OLER RAPE SUNF SUTA PULS POTA TEXT TOBA TOM OVEG APPL OFRU CI		dummyTrd	set	1	1		
		META_STEP	set	1	1		
000000	<u> </u>	countriesMeta	set	1	35		
IR		META	set	4	16401		
BA		SmulationYear	set	1	20		
ED		othi	Set Company				
20PAR	=	step	set	1	99		
ER OTH		SCEN	set	1	2		
		VEADO	***		٨		
IS 0.00		T:\	britz\capri\gams\convert n	narket mod	el 1.qdx		
GR	II.	er input name	type	dims reo	ords k	ong text	
	-	lines	set	1	10001	A	
NZ		dummyTrd	set	1	1		
30		META_STEP	set	1	1		
OR		countriesMeta	set	1	35		
IDEAST		META	set	4	16401		
3A		Simulation tear	set	-	1		
H		ntelul	narameter		1		
NF		step	set	1	99		
FR LDC	-	SCEN	set	1	2		
e m		VEADC		-	A		
quations	<ul> <li>Variables</li> </ul>						
2Val_(NO000000,OFRU) "Value of imports "	pv_bevFuncIntAc	ldFac		.1=1	59,374	.lo=0 .upp=1.5	593.7
9.45232123664149e-7*v_armlPrice(EU010000, POTA)*v_armlQuant(EU015000, OLIO)	v_armlPrice (EUG	010000, POTA)	Armington first stage pr	.1=1	.366,5	38.lo=136,654.upp=13.	665,
-9.45232123664149e-/*(v_entryFrideDiver(LOUISOUD,BRA,CITR)*v_feedInpCoeff			Armington first stage				
+v_intervStockChange(EU015000,BEEF)*pv_bevFuncIntAddFac	E warm10uant (FII)	15000 0170)	aggregate, beware: uilit	y 1=7	73 444	lo=0 upp=773	2 4 4 4
*v_impFrice(N0000000, MED, CITR) *v_tradeFlows(N0000000, MED, CITR)	- analyuane (Eoc		points, in calibration p [1000 tons]	oint .1-/		.10 0 .upp-//3	
<pre></pre>	v_entryPriceDri	Lver (EU015000, BRA, CITH	Reduction of specific ta	.1=1	.313,6	61 .lo=131,366 .upp=131	1.366
Quant_(USA,OFRU) "Sum of exports of each region "	v expOuant (ASO	CE REST, RYEM)	Sun of exports [1000 ton	.1=3	48,142	.upp=7.1	118.4
	v feedInpCoeff		Feeding per head and yea	r in .1=6	0.538	.lo=0 .upp=605	5.384
-0.0028/15631698088/*pv_bevFuncIntAddFac			kg			impose	
-0.00287156316980887*v_tradeFlows(TUR,RUS,SUNO)			Import price for goods f BM1 in BM after border.	rom			
-0.00287156316980887*v_tradeFlows(MER_0TH,EU015000,WMIO)	v_impPrice (NOO)	00000, MED, CITR)	including tariffs, [curr	.1=1	.231,9	77.lo=123,198.upp=123	3.197,
-0.00287156316980887*v_tradeFlows(CH,EU015000,BUTT)			Euro/ton]				
-0.00287156316980887*v_tradeFlows(IND,AFR_LDC,PULS)			Import price for goods f	zom			
-0.00287156316980887*v tradeFlows(CHN, THAI, OCER)	v_impPrice (NOO)	00000,RUS,RAPO)	RM1 in RM after border, including tariffs, lourn	.1=2	.554,0	18 .lo=255,402 .upp=255	5.401,
	+		Euro/ton]				

In that case, it will replace working with the output of solprint=1 in combination with limrow/limcol. The necessary output can be produced as seen in the example below from "arm\prep\_market.gms" in the CAPRI code:



The different files which can be used with the viewer are described briefly below:



#### Includes

When the viewer is started, only the code for selected GAMS file is loaded in the window titled "GAMS". The window titled "GAMS includes" will show all the (direct) includes used by that GAMS file which could be successfully read (globals are not yet treated, it is planned to merge the tool with the HTML doc generation such that includes reflect the compilation stage of the GAMS code.).

In order to open included files:

- either double click on a node in the "GAMS includes" tree view
- or on a "\$(bat)include" in the GAMS code view marked in red

As seen below, the GAMS code in new file will be opened in an additional tab. In case the new file comprises includes, they will be added to its node.



#### Loading symbols

The GAMS code highlights in blue all symbols (sets, parameter, variables, and equations) found in the GDX file. These symbols can be opened in the GDX viewer by a mouse click in the GAMS file viewer. If several GDX files are provided, the symbol will be loaded from all the GDX files where there are non-empty records. That allows for a very rapid inspection of the data.

	4	GDXV	iewer
	file	Vie	ws
	1	5	Files
	ſ	-	Selection
			Gams code
	ш		Equations
			Variables
	ш		Gams includes
(CDV) = 1 + 1 + 1 + 2 + 2 + 2 + 4 + 3 + 3 + 2 + 2 + 2 + 2 + 2 + 2 + 2 + 2	Ľ		GDX symbol table
"GDX symbol table" from the "Views" ment	2	4	GDX symbol view

Alternatively, select the

i tubics iouu	ea from GDX file(s)					
	T:\brit	z\capri\gams\convert	_market_	model_2.gd	x	
user input	name	type	dims	records	long text	
	lines	set	1	10001		
	dummyTrd	set	1	1		
	META_STEP	set	1	1		
	countriesMeta	set	1	35		
	META	set	4	16401		
	SimulationYear	set	1	1		
	Countries	set	1	35		
	ntslvl	parameter	0	1		
	step	set	1	99		
	SCEN	set	1	2		
•	VEADS	ant	,			•
•	T:\brit	z\capri\gams\convert	_market_	model_1.gd	x	4
< user input	T:\brita	z\capri\gams\convert	_market_	model_1.gd records	x long text	ŀ
د المعالم ( user input	T:\britz	z\capri\gams\convert type set	_market_ dims	model_1.gd records 10001	x long text	4
د ا user input	T:\britz	z\capri\gams\convert type set set	_market_ dims	model_1.gd records 10001	x long text	4
< user input	T:\britz Iname Iines dummyTrd META_STEP	z\capri\gams\convert type set set set set	_market_ dims	model_1.gd records 10001 1 1	x long text	4
< user input	T:\britz name lines dumyTrd META_STEP countriesMeta	z\capri\gams\convert type set set set set set	market_	model_1.gd records 10001 1 1 35	x long text	4
< user input	T:\brit. name lines dummyTrd META_STEP countriesMeta META	z\capri\gams\convert type set set set set set set	_market_ dims	model_1.gd records 10001 1 1 35 16401	x long text	•
< user input	T:\brit: Ines Ines dumnyTrd META_STEP countriesMeta META_SimulatonYear	ctopri\gams\convert type set set set set set set set	_market	model_1.gd records 10001 1 1 35 16401 1	x long text	4
< user input	T:\britz name ines dummyTrd META_STEP countries SimulationYear Countries	ct caprilgams convert type set set set set set set set set	_market_ dims	model_1.gd records 10001 1 1 355 16401 1 355	x long text	4
< user input	T:\britz name Ines IdumyTrd META_STEP CountriesMeta META SimulationYear Countries ntsW	set set set set set set set set set set	_market dims 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	model_1.gd records 10001 1 1 355 16401 1 355 16401 1 355 1	x long text	•
vser input	T:\britz name Innes dummyTrd META_STEP countresMeta META SimulationYear Countries ntsM step	set set set set set set set set set set	_market	model_1.gd records 10001 1 1 355 16401 1 355 16401 1 355 16401 99	X long text	•

In table above, symbols can be selected by the mouse. As in the GDXViewer, one might select several symbols with the same number of dimensions.

#### Working with the equation and variable viewer

Equations and variables can be loaded in the equation and variable viewers (the two windows in the lower part of the main window) by working with the selection boxes (see tab selection). These windows are thought as a replacement of inspecting the equation listings with an editor.

If an equation is selected with the **"equation selection" control**, the equation window will show it in "linearized" form, i.e. any non-linear functions and interactions terms with other variables are converted in a constant. The variable window will report the level and the lower and upper bound for all variables found in the equation. The user can add (or remove) equations with the control as well.

If a variable is selected with the **"variable selection" control**, the variable window will report the level and the lower and upper bound for the selected variable(s). The user can add (or remove) variables with the control as well.

Double clicking on a variable in the equation window will show that variable in the variable window as well as load it in the GDX viewer.

Double clicking on a variable in the variable window will load all the equations comprising the variable, and also show in the variable window all the variables found in these equations.

The first line of selection boxes shown below first allows to select all instances of a variable or an equation:

Variables v_GLDemandFS 🗸 🗸	Equations ArmBal1_	1* •	2 *	→ 3 <sup>*</sup> →
----------------------------	--------------------	------	-----	--------------------

The other boxes show all items found on the dimension of all symbols and allow to filter further. In order to see e.g. all equations which have on the first dimension "AFR\_LDC" and on the second "APPL", put the selection controls as seen below:

Equations \* 1 AFR\_LDC V 2 APPL V 3\* V

That produces an equation output as seen below:

```
    Equations
    ArmBall_(AFR_LDC, APPL) "Adding up of human consumption feed and processing " ..
    -0.0114645139135529*v_consQuant(AFR_LDC, APPL)
    +0.0114645139135529*v_armlQuant(AFR_LDC, APPL) =E=0;
    CPri_(AFR_LDC, APPL) "Consumer price linked with fixed margin to first stage price " ..
    -0.00187949947079955*v_armlPrice(AFR_LDC, APPL)
    +0.00187949947079955*v_consPrice(AFR_LDC, APPL)
    +0.00187949947079955*v_consPrice(AFR_LDC, APPL) =E=0.284525356896553;
    GLDemandGiS_(AFR_LDC, APPL) "Part of GL demand function " ..
    -0.521399433800188*(0.508934628613077
    +1.14603816746872*sqrt(0.00188303864092867*v_consPrice(AFR_LDC, MAIZ))
    +1.37181255843937*sqrt(0.00188303864092867*v_consPrice(AFR_LDC, OCER))
```

Note: The maximal number of equations can be changed in the top panel.

## Decomposition

If a variable on the LHS depends on several variables on the RHS, it might not be easy to understand what drives the change of the LHS. Examples of such relations provide behavioral equations with own- and cross price effects or profit changes depending on changes of multiple netput quantities and prices. Understanding these changes is often crucial for policy analysis (and during debugging). The equation and variable viewer therefore provides a decomposition of changes on the LHS of the equations, based on the Jacobian provided by GAMS and the simulated changes of the variables against their starting point.

That requires that the CONVERTD solver from GAMS produces the information on the Jacobian as shown in the example below, and that information is loaded into the equation and variable viewer:

\$iftheni.after "%1"=="after"	
<pre>\$onecho &gt; convertd.opt</pre>	Produces a GDX
gams %gamsDocDir%\convert_gtap8.gms dict %gamsDocDir%\convert_gtap8.txt jacobian %gamsDocDir%\convert_gtap8_after.gdx	with the variable levels after the solve,
\$offecho	the Jacobian, and
gtap.optfile = 1;	a linearized version
\$else.after	of your model as GAMS code
\$onecho > convertd.op2	
jacobian %gamsDocDir%\convert_gtap8_before.gdx	
\$offecho	
gtap.optfile = 2;	Produces a GDX
\$endif.after	with the variable
<pre>gtap.holdfixed = 0; option nlp=convertd; solve gtap using nlp minimizing dummy ; gtap.holdfixed = 1;</pre>	
); Sendif CONUERT	
	Call to CONVERT,
	Holdfixed = 0 allows to track
	effects of fixed variables

Based on the information, the viewer will try to detect the variables in the LHS. If Jacobian information is available, the viewer will provide additionally the following information:



Formally, the calculation of the estimated contribution of a RHS variable to the changes on the LHS can be described as follow:

Change in LHS variable x approximately equal to changes in RHS variables y weighted with their Jacobian entries

CONVERT utility from GAMS x = f(y)delivers Jacobian  $\Delta x \approx \sum$ CONVERT utility from GAMS delivers variable levels before and after the solve to calculate differences in x and y

And:

In order to easier digest the information, contribution of variables y is expressed as % contribution to change in x, entries add up approximately to 100%



## Utilities: Gdx-file(s) viewer

GDX-files are generated by GAMS and typically serve either an exchange format between different GAMS applications, or for exploitation purposes as the GAMS-IDE comprises a view for GDX-files. Further tools for GDX-files are available from GAMS company and are described in different documents. In opposite to listings generated by GAMS programs, the GDX files store the data in full numerical precision in an internal format.

The new CAPRI version passes information from one task to the next with the help of GDX files, so generates CoCo a gdx files with the time series at national level, which is read by CAPREG. And the regional time series generated by CAPREG are inputted by the trend projection tool CAPTRD. These gdx files are accessed when the different tasks of "Data base exploitation" are chosen. The user has on top the possibility to load one or several tables from one or several freely chosen gdx files.

The GDX exploitation utility can be reached via the menu bar:

(	Utilities GUI Help
5	GDX Viewer
ē	Start equation and variable viewer
•	Build HTML documentation
4	SVN update
	Use 't:\britz\capri\gams\util\cleanUpGamsDir.bat' to clean up source directory

#### Graph: Panel to GDX file exploitation

Loa	ad gdx file	Load sel	ected tables(	5)	Set XML tab	finition file	s from : tables.xml	
List of tables loaded from GDX	t of tables loaded from GDX file(s)							
			T:\britz\	CAPRI\r	esults\Capre	ert_out.gdx close		
	user input	name	type	dims	records	ig text		
		fert_out	parameter		2 9740			
			T:\britz\C	APRI\re	sults\Capreg	t_out_F.gdx Close		
	user input	name	type	dims	records	ig text		
		fert_out	parameter		2 9768			

When the task "exploit gdx files" is selected by pressing the related button, four buttons are shown in the task panel. The first one, labelled "load gdx files" will open a file selection menu when pressed. When the ok button of the dialogue is operated, the content of the gdx file is partially loaded, and a table is added to the right upper window of the application showing the parameters and sets comprised in the gdx files, along with their number of dimensions and records. When the close button next to the table is pressed, the table is deleted. Pressing the "load gdx file" again will add more tables.

One parameter from each table may be selected (pressing the "crtl" key when clicking with the mouse de-selects). If several parameters from one file need to be loaded, the user may open the same file several time. The content of the different parameters is merged together, and the parameters themselves span an additional data dimension. If the user does not provide input in the first column of the tables labelled "user input", the program will generate names automatically. The data loaded are shown in the table tool described above.

The user can use view definitions stored in a XML file to the tables by pressing the enabling the "Use table definitions from ..." tick box, and may use the "Set XML table definition file" button to change the file to use.

# Utilities: Generating coordinate files for the exploitations tools from shapefiles

The exploitations tools use a proprietary format to store coordinate files. The utility allows to build from shape files a file in that proprietary format.

GL	II Settings Help				
	Batch execution				
	Generate GUI geometry from	shapefile			
e ha	seline				
<u></u>					
Coc	o file generation from shape file	) T:\britz\capri	\addon\gis\capri.shp	Set file	
	ShapeFile (inpu	) T:\britz\capri	\addon\gis\cities Project.sh	Set file	

ShapeFile (input) T:\britz\capri\addon\gis\cities_Project.shp	Set file
ShapeFile (input) T:\britz\capri\addon\gis\main_rivers.shp	Set file
ShapeFile (input)	Set file
ShapeFile (input)	Set file
ShapeFile (input)	Set file
Coordinate file (output) T:\britz\capri\GUI\test.zip	Set file
Id string NUTSII 🗸 Name string NUTSII 🗸 Scaling for coordinates 0.001 Fill up mask	for ids
Id string NAME 👻 Name string NAME 👻 Scaling for coordinates 0.001 Fill up mas	sk for ids
Id string INT_NAME  Vame string INT_NAME  Scaling for coordinates 0.001 Fill up ma	isk for ids
Id string 💽 Name string 💽 Scaling for coordinates 1. Fill up mask for ids	
Id string 🔍 Name string 🔍 Scaling for coordinates 1. Fill up mask for ids	
Id string 🔍 Name string 🔍 Scaling for coordinates 1. Fill up mask for ids	
Simplify Store holes Minimum size for polygon 50	
PFAFSTETTE 1.0 3	<b>^</b>
MAINDRAIN_ 5000 4	
INT_NAME Oka, Volga 4	
WINDOW 2013 5	
MAIN_PER_W y 5	
OBJECTID 1 6	=
SHAPE_Leng 3710847.47957 7	-
	•
Get fields from cooFile Start coordinate generation	

**Note:** Files and settings shown above are the ones used to generate the NUTS II map in CAPRI, the shape files can be found in the "addon\gis" folder

Coo file generation from shape file	۱ ۱
ShapeFile (input) T:\britz\capri\addon\gis\capri.shp	Set file
ShapeFile (input) T:\britz\capri\addon\gis\cities_Project.shp	Set file
ShapeFile(input) T:\britz\capri\addon\gis\main_rivers.shp	Set file
ShapeFile (input)	Set file
ShapeFile (input)	Set file
ShapeFile (input)	Set file
Coordinate file (output) T:\britz\capri\GUI\test.zip	Set file
Id string NJTSII  Vame string NJTSII  VScaling for coordinates 0.001 Fill up mask	for ids
Id string NAME  Vame string NAME  Vame string NAME  Vame string Scaling for coordinates 0.001 Fill up mass	sk for ids
Id string INT_NAME  V Name string INT_NAME  V Scaling for coordinates 0.001 Fill up ma	isk for ids
Id string 🔍 Name string 🔍 Scaling for coordinates 1. Fill up mask for ids	
Id string 💽 Name string 💽 Scaling for coordinates 1. Fill up mask for ids	
Id string 🔍 Name string 🔍 Scaling for coordinates 1. Fill up mask for ids	
Simplify Store holes Minimum size for polygon 50	
PFAFSTETTE 1.0 3	A
MAINDRAIN_ 5000 4	
INT_NAME Oka, Volga 4	
WINDOW 2013 5	
MAIN_PER_W y 5	
OBJECTID 1 6	=
SHAPE_Leng 3710847.47957 7	
	•
Get fields from cooFile Start coordinate generation	

As a first step, the shapefiles must be analyzed by using the "Get field from coofile" button:

Once, that is done, the fields from the shapefiles used for keys and the long texts can be chosen, and some other settings:

- **Simplify**: will simplify the geometry by shifting points, using a pre-defined distance in the coordinate room. That distance is current hard-coded and matches data sets in universal meters.
- Store holes: lake and similar features which are not part of a polygon are drawn by the mapping tool in while. Attention: that should not be used if regions are "island" inside other regions.
- **Minimum size of polygons**: will drop polygons below the specified threshold. Try 0 first to avoid that no output is generated.

The interface will assume treat line strings as river, points as cities and polygons as regions.

The utility assumes that all shapefiles are in the same coordinate system and will simply store the coordinates one to one (applying where set a scaling factor) in the internal format used by GGIG. It generates in parallel to the coordinates a file called with the same name as the outfile file with the extension zip, in the example above, that file would be called "test.gms". it comprises the list of regions ids read from the shapefile and a GAMS parameter which reports any numerical fields found in the shapefile. If no numerical fields are found, the position of the regions are stored in the parameter. The file has accordingly the structure shown below:

```
SET rr /
AT11 'AT11'
AT12 'AT12'
AT13 'AT13'
AT21 'AT21'
/
PARAMETER TEST /
"AT11"."area_ha" 396553.0
"AT11"."Center" 4811506.3703
"AT11"."Center" 4811506.3703
"AT12"."area_ha" 1919280.0
"AT12"."Center" 4747718.75342
/;
execute_unload 'test.gdx' TEST;
```

When it is executed, it generates a GDX file which can be loaded with the GDX Viewer

🛃 GDX viewer		
View Handling Windows		
🕼 TEST (T:\britz\capri\GUI\test.gdx) [0]		
<u> </u>	area_ha	Center
<b>T</b>		
AT11	396553.00	4811506.50
AT12	1919280.00	4747719.00
	41472 00	4794673.00

After a switch to mapping view, and the newly generated coordinate set can be loaded and checked:



Clearly, that proceeding allows inspecting numerical features in a shapefile without requiring a GIS program.

## Analysis differences in GAMS based data using GGIG

## Background

In result analysis such as when comparing scenarios, but also when comparing different releases of data sets against each other, one frequently wants to see only those values with larger changes. When working with values generated by GAMS, one has different ways to proceed.

If the two data sets to compare are comprised in GDX files, one can use the GDXDIFF utility from GAMS. The dis-advantage is that GDXDIFF does not have information about the logical structure of the data or its content, such that it might be cumbersome to filter our large absolute or relative changes which matter.

The other extreme is to write a GAMS program which uses rules about the importance of data items to concentrate on changes which matter. In the CAPRI-RD project, to give an example, outlier statistics for many time series for EU Member States were calculated, leading to thousands of potentially "suspicious" values. It is clearly impossible to check manually each and every case, so that algorithms have to deal with the majority of the cases. The expensive manual checks have to concentrate on the items which are deemed important. Hence, by using e.g. national and EU crop shares, animal stocking densities and shares on sectoral revenues, a matrix of importance was constructed which assign a numeric indicator to each time series. A combination of that importance metric for a time series and its outlier statistics combined with a threshold delivers then the potential outliers to compare manually.

The following will describe a third way to proceed based on the in-built functionalities of the exploitation tools.

## Comparing two data sets in GGIG, example from CAPRI

Generally, each task in CAPRI also allows viewing its results and selecting "Scenarios". The example below show the task "Build regional time series". On the disk, several versions where located, and these can be compared as if they were scenario. If one e.g. wants to compare the current version against the trunk or an earlier release, one can rename the current one .e.g. to "…\_current.gdx" and then use "update" or "update to revision" to download the version to compare to from the server.

There are now different way how to proceed.

#### **GGIG as GDXDIFF**

One can mimic the behavior of GDXDIFF by using as the view "No table"

iew h	landling Windows		
£ Su	pply details [0]		
	Scenario shifter	*	
	Welfare	٠	
	Markets	*	۰.
1	BioFuels		
U	Prices		
-	Farm		
9	Farm - totals	,	
0	Supply model analysis		
-	Farm EU	,	
0	Fertilzer		
v	Feed		
P	Environment		
"	Energy		
s	HSMU	,	
la	DNDC		
^	ESIM comparison		
	No table		

That allows scrolling through any cell, but requires familiarity with the codes and the structure of the data set one analyzes. One might end up with a view as below ... the first to do is to pivot.

🛃 No tal	B No table [0]													
	Region				Year									
	КО000000				<ul> <li>RES_TIME_SERIES</li> </ul>	RES_TIME_SERIES								
		1984												
			1	1	1									
		BIOE	BIOD	SWHE	DWHE	RYEM	BARL	OATS	MAIZ	OCER	RAPE	SUNF		
¥														
SWHE														
DWHE														
RYEM														
BARL														
OATS														
MAIZ														
OCER														
RAPE		_												
SUNF		_												
SOYA														
OOIL		-												
OIND		-												
NURS		-												
FLOW		-												
MATE		-												
ROOF		-												
OFAR														
GRAF														
GRAI														
PARI														
OLIV														
PULS														
ΡΟΤΑ														
SUGB														
TEXT														
TOBA														
TOMA														
OVEG														
APPL		_												
OFRU		_												
CITR		_												
TAGR		_												
TABO		_												
TWIN														

Imagine you want to check acreages across the regions. A good way to proceed is to put the regions in the rows and the years in the columns as shown below.

STransposing and Me	erging	
Table control area		0
Box (1)		250
Product(250)		
Box (2)		2
versions(2)		
Box (3)		205
Activity(205)		
lable area	26	Table column groups Table columns years(26)
Table row groups	Table rows Region(387)	Table cells area
	ok	Cancel

Next, one uses the option dialogue to configure the view such that e.g. percentage differences against the old version are displayed:

Sustomize view
Tahoma 🔹 🚺 🔹 plain 💌
Fraction digits and decimal separator
Separator between merged data dimensions
Column width 140 📩 Row width 140 🖈
Hide empty rows Hide empty columns
Cut off limit to determine empty cells
Vuse default pivoting for tables 🔲 Show histogram 👘 Use classification colors for tables
Show only selected items
Comparison output Values and percentage difference 👻
Comparison threshold to hide values 0
Region Activity Data dimensions used for comparisons Product years versions
Element used for comparisons versions
RES_IME_SERIES RES_TIME_SERIES161111
ok define colors define statistics store settings load settings

That produces a view as seen below ... but it is clearly not inviting to scroll now through about thirty years and almost 400 regions.

📕 No ta	ible [0]											
	Product			versions			Activity		Percentag	e diff. to		B
	LEVL			RES_TIME_SERIES		•	<ul> <li>SWHE</li> </ul>		RES_TI	ME_SERIES161111		벽
	1	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
¥												
коооо	000											
BLOOO	000	191.82 1.23%	2 193.53 6 1.33%	193.94 0.19%	198.26 1.43%	217.70 -1.25%	219.46 0.03%	218.99 -0.40%	213.9 -0.32	3 214.54 % -0.509	4 211.95 6 0.98%	
DK000	000	318.33 0.67%	3 326.64 6 0.991	341.43 5.34%	381.43 1.02%	296.88 2.91%	434.96 2.03%	516.55 1.01%	5 504.1 1.27	0 559.60 % 3.45%	0 601.09 6 2.03%	1
DE000	000						2383.88 0.00%	2383.88 0.00%	8 2474.0 0.93	14 2621.60 % 1.95%	5 2408.91 6 1.51%	24
ELOOOC	100	506.60 -4.67%	0 438.80 6 0.87%	387.80 -3.15%	369.70 -3.52%	357.11 -2.96%	350.51	309.33 -1.50%	280.3 -3.17	4 309.78 % -2.96%	8 307.67 6 -1.16%	1
E50000	000	2038.83	3 1817.29 6 -1.49%	1898.56 -1.48%	2012.36 -1.49%	2163.35	2100.96 -1.00%	1736.56 -0.51%	5 <b>1686</b> . -2.48	7 1526.07 % -3.86%	7 1338.84 6 -2.42%	15
FR000	000	5025.58 1.51%	8 4695.08 5 1.86%	4600.89 3.71%	4611.01 4.41%	4500.03 5.12%	4820.48 3.05%	4868.61 1.66%	4770.0 2.53	13 4769.83 % 2.02%	3 4440.35 6 3.00%	44
IR0000	000	72.19	9 70.65 6 -5.35%	72.75 -2.96%	56.38 -2.46%	58.60 -3.43%	60.70 -2.54%	-0.50%	86. 0.00	6 91.00 % -0.94%	5 79.20 6 0.50%	
110000	100	1375.60 11.949	0 1213.49 6 -1.21%	1190.63	1133.53	1050.04	1104.07	1009.22	953.	0 963.35 % 0.509	5 872.49 6 -0.88%	

Here, three combined options in the dialogue can help:

(1) Use hide empty and empty columns to throw out missing values or hidden cells.

(2) Use a (approximate) cut-off for the value to show, e.g. start only with acreage > 1000 [1000 ha]

(3) Hide cells where the difference is below a threshold, e.g. 1%

🔹 Customize view							
Tahoma 🗸 11 🗸 plain 👻							
Fraction digits and decimal separator 2 🔹 🗸 💌							
Separator between merged data dimensions							
Column width 140 📩 Row width 140 🖈							
Hide empty rows							
Cut off limit to determine empty cells 1000,00 💌							
☑ Use default pivoting for tables							
Show only selected items							
Comparison output Values and percentage difference							
Comparison threshold to hide values							
Region Activity Data dimensions used for comparisons Product years versions							
ok define colors define statistics store settings load settings							

That delivers a much more usable view at important changes:

No table [0]												E ×
Product			versions			Activity		Hide <= 1.	0, Percentage diff. to		D	View type
LEVL			RES_TIME_SERIES			SWHE		* RES_TIM	E_SERIES161111	Pres-	140 88°	Table
8	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993 199	4 199	5
TE000000	-								2621.6	5 2408.91		
E500000	2038.8	1817.2	9 1898.5	i6 2012.36				1686.5	7 1526.07	1338.84	1297.10 -2.86%	1415.56
FR000000	5025.50 1.519	4695.0	8 4600.8 4 3.71	19 4611.01 % 4.41%	4500.03 5.12%	4820.4	8 4868.6 5 1.469	4770.0	3 4769.8 6 2.02%	4440.35	4467.07 2.97%	4627.43
11000000	1375.60 11.949	1213.4	9 1190.6 • -2.88	3 %							819.87 -1.49%	
UK000000	1889.6 46.001	1792.2 55.201	9 <b>1923.5</b> 4 46.79	6 1808.79 % 49.30%	1794.32	1998.5 28.209	6 2070.9 6 9.25%	2080.6	7 2071.54 6 21.929	5 1877.34 6 6.31%	1994.24 34.08%	2076.13
HU000000								1081.6	5 889.00 6 -1.30%	0 1029.34 -1.69%		
PL000000							2169.20 1.929				2289.32 1.01%	
8600000							1106.1-		1027.75			1117.88
R0000000						2239.1 1.529	1 2333.2 6 4.539	2122.1	5 146.20 6 2.029	5 2236.42 2.02%	2392.37 3.09%	2431.68 2.02%
TUR												7529.37
ES400000	1292.0 -6.939	1151.5	9 1199.1 6 -6.12	15 1272.41 % -6.20%	1371.26	1329.3	8 1095.5 6 -5.049	1065.4 -4.59	8 953.81 6 -7.709	835.38 -4.42%	795.28 -6.77%	876.17 -6.32%
FR200000	2527.8 3.219	2357.6 3.819	2 2323.6 5.04	7 2329.36 % 5.42%	2269.87 6.31%	2427.3 4.789	5 2407.8 6 3.75%	2368.9	5 2333.65 4 4.00	2244.84 4.78%	2270.75 4.77%	2329.32 5.57%
FR500000			830.4	14 832.52 % 7.89%	821.23 7.93%	894.4	1 898.3 6 2.639	900.8 3.23	8 918.14 6 2.849	869.51 3.61%	896.80 3.47%	891.38
TR600000												
TR700000												1057.54 -2.65%

An interesting option here is to use the "GTAP" difference which is defined as

 $m = \left[\log(x) - \log(y)\right]abs(x - y)$ 

The first term looks at relative differences which are weighted with absolute ones.

#### Using the table definitions

For those not familiar with the codes, it might be easier to work with the pre-defined tables. As those tables are not always constructed e.g. to be used with time series, it might be required to pivot them as well.

🗟 Supply details (0)														
Region versions					Item						Hide <= 5.0, Percentage diff. to			
RES_TI			RES_TIME_SERIES		-	Hectares or herd size [1000 ha or hds]				-	RES_TIME_SERIES161111			
-	<b>Y</b>	1984	1985	1986	1987	1988		1989	1990	1991	1997	L. C.	1993	1994
¥													1	
Utilized	l agricultural													
Oilseed	ls									1073.71		1489.72	1752.2	5
Other a	arable crops	1094.70 7.15%	1045.33	1036.32	1029.74		1062.57	912.60 8.03%	927.55	918.26 12.05%		838.66	725.8	1
Fodder	activities												9275.66 -5.069	0
Set asid	de and fallow											3560.38 -10.78%		
All cattl	le activities	3524.90 -27.01%	3699.85 -29.77%	3676.01 -33.48%	3692.47 -36.58%		3725.92 -37.17%	3786.42 -37.11%	3808.13 -38.48%	3845.82 -39.29%		3886.38	3759.79	9
Beef me	eat activities	607.87 -42.93%	752.53 -43.81%	<b>\$18.96</b> -49.50%	928.83 -53.63%		1007.36 -50.74%	1049.77 -50.15%	1116.77 -51.97%	1210.29 -51.53%		1320.02 -50.86%	1453.83 -47.699	3 6
Other a	nimals	4805.70 -90.36%	4598.54 -90.33%	4774.46	5476.63 -89.94%		6087.76 -90.05%	6257.88 -89.69%	6464.73 -89.77%	6691.29 -89.59%		6946.60 -89.28%	7028.10 -89.679	6
Utilized area	l agricultural													
Oilseed	ls									1073.71 -6.26%		1489.72 7.24%	1752.25 11.599	5
Rape										11.96			16.8	1

## Comparing two GDX files with GGIG

<u></u>	APRI [t:\britz\capri\gams]					
File	Utilities GUI Settings Help					
CA	GDX Viewer					
٢	CSV to GDX					
0	Start Command prompt					
$\odot$	Start equation and variable viewer					
CA	Build HTML documentation					
$\odot$	SVN update					
0	Use 't:\britz\capri\gams\util\cleanUpGamsDir.bat' to clean up source directory					
◎ Find No of fts in Nuts2						
$\odot$	Define fts from FSS					

GGIG also comprises a GDX viewer:

Into which several GDX files can be loaded and compared as discussed above.

## Index

Batch execution 135	Drop-down boxes for selections 33
Clipboard export 43	Equation and variable viewer 155
Column and row selection 35	Export to file 43
Data export 43	GDX viewer 163
Data export to ArcGIS 105	Generate GAMS documentation 149
Data export to EXCEL 47	Generating coordinate files 165
Decomposition 162	Graphics 55

Bar charts 62	Equal interval 88
Box and whisker charts 68	Excluding zeros from classification 86
Clipboard export 60	Manual classification 88
Cummulative distribution 70	Mean standard dev 88
Deviation renderer 73	Natural breaks 87
Export to file 59	Nested mean 88
Histograms 69	Quantile 87
Line chart with mean/min/max 74	Clipboard export 103
Line, point and area charts 64	Coordinates 83
Markov charts 77	Drag 95
Network charts 76	File export 103
Pie charts 66	Flow maps 78
Scatter plots 71	Frequency diagram in map 89
Setting colors 58	Full extent 95
Spider plots 67	Getting data for specific polygons 96
Waterfall charts 63	Highlighting specific regions in the map 97
Help system 26	Histogram window 84
Histogram 41	Info pointer 96
legend	Info pointer and window 96
continuous linear scaling 93	Legend 93
continuous logarithmic scaling bar 94	Pie chart maps 80
Machine learning 112	Pop-up menu 105
Maps	Regional labels in map 101
Classification	Rivers and cities 102
Area weighted classification 86	Schemes 104
Classification method 86	Shrinking polygons according to UAA share
Color for middle class 92	85
Color table 89	Store settings 103
	Title 95

Updating the map 101	Result directory 12
Zoom in 95	Set up
Zoom out 95	Look and feel 13
Zooming 95	Starting GAMS 26
Menubar	Tables
File menu 134	Drill down 43
Settings menu 135	Filtering 49
Utilities and GUI menu 135	Outlier detection 52
Meta data 131	Pop-up menu 54, 60
Concept 131	Sorting 48
SET META 132	Statistics 49
SVN information 134	View options 38
Numeric filter 98	Classification colors for tables 40
Pie chart maps 81	Column and row width 39
Pivoting 37	Comparison output 40
Predefined selection groups 36	Cut off limit to determine empty cells 39
Scenario editor 122	Fonts 39
Select scenarios 29	Hiding empty columns or rows 39
Set up	Histogram window 40
Exploitation tools 20	Number formatting and rounding 39
GAMS.EXE 14	Percentage differences 40
SVN 15	Short codes and/or long texts 40
Work directory 12	View Selection 32
Set up 12	View type selection (tables, graphs, maps) 36