Analyzing Results from Agricultural Large-scale Economic Simulation Models: Recent Progress and the Way Ahead

Gegenwärtige Entwicklung und Perspektiven bei der Analyse von Ergebnissen komplexer ökonomischer Simulationsmodelle

Wolfgang Britz
University Bonn, Germany

Ignacio Pérez-Dominguez
European Commission, Seville, Spain

Gopalakrishnan Badri Narayanan
Purdue University, West Lafayette, IN, USA

Abstract
This article discusses the role of post model processing and Graphical User Interfaces (GUIs) in the context of applying large and complex economic simulation models for Impact Assessment (IA) of policies impacting the agricultural and food sectors. The calculation of indicators from different domains and decomposition approaches is presented as part of post-processing of model results. We review user roles with regard to the application of complex economic models in agricultural IA, concluding that the GUIs will be mostly used by experienced analysts. The article stresses the role of GUIs in separating the code implementation of the economic model from the presentation of their results for analytical purposes. It highlights core functionalities of GUIs based on a detailed comparison between the GUIs and post-model analysis tools of three large economic models: Aglink-Cosimo, CAPRI and GTAP (Global Trade Analysis Project). Whereas the Aglink-Cosimo and CAPRI (Common Agricultural Policy Regionalised Impact Modelling System) models follow a strict separation between model and data analysis software tools, the GTAP model follows a model integrated approach using GEMPACK (General Equilibrium Modelling Package) tools. Moreover, the CAPRI GUI is technically and formally detached from the data generation process. Aglink-Cosimo and GTAP models use tools specifically developed to serve mainly their own needs (and tailored to communicate with their model drivers). One specific characteristic of Aglink-Cosimo relates to the need of interactive documentation for users due to the declarative nature of the code, which relies upon the obligation of the OECD and FAO to deliver a transparent and well-documented model and database to a formalized group of users. The CAPRI and GTAP models are much more research-driven and therefore have invested more resources in building bridges with other model consortia. We conclude that an integrated GUI would benefit from combining the best features of the different approaches presented.

Key Words
economic modelling; policy impact assessment; post-model processing; graphical user interface, agriculture

Zusammenfassung
nissen der folgenden ökonomischen Modelle: Aglink-Cosimo, CAPRI (Common Agricultural Policy Regionalised Impact Modelling System) and GTAP (Global Trade Analysis Project). Aglink-Cosimo and CAPRI wenden jeweils unterschiedliche Software für Simulationen und Ergebnisanalyse an, während GTAP einen integrierten Ansatz basierend auf GEMPACK (General Equilibrium Modelling Package) verfolgt. Die CAPRI GUI abstrahiert darüber hinaus technisch und auch formal vom spezifischen Datengenerationsprozess des ökonomischen Modells. Im Gegensatz hierzu sind die Analysertools von Aglink-Cosimo und GTAP weitestgehend spezifisch für diese Modelle entwickelt worden (und technisch auf die jeweiligen Modelltreiben ausgelegt). Aglink-Cosimo bietet als spezifische Charakteristik eine interaktive Dokumentation an. Dies reflektiert einerseits die deklarative Natur des Modell-codes und andererseits die Verpflichtung seitens der OECD und FAO, einer formal definierten Nutzergruppe Modell und Datenbank transparent und wohl dokumentiert zur Verfügung zu stellen. CAPRI und GTAP sind hingegen in höherem Grade in Forschungsaktivitäten eingebunden, was sich auch in den zahlreichen Schnittstellen mit anderen Modellsystemen wiederspiegelt. Der Beitrag folgt, dass eine modellübergreifend einsetzbare GUI, die die Kern-eigenschaften der verschiedenen untersuchten Ansätze integriert, von Vorteil wäre.

Schlüsselwörter
ökonomische Modellierung; Politikfolgenabschätzung; Aufarbeitung von Modellergebnissen; graphische Benutzeroberfläche, Landwirtschaft

1 Introduction

Governmental institutions (e.g. EUROPEAN COMMISSION, 2009) frequently request formal Impact Assessments (IA) before launching legislative proposals; in the case of agricultural policies, these are often partly based on economic simulation models (cf. BRITZ et al., 2012; THIEL, 2009). However, many tools especially developed for IA, such as those by the EU Research Framework Programs seem to be under-utilized (NILSSON et al., 2008; JANSEN et al., 2012). This is a critique also raised by governmental bodies themselves, e.g. by the European Commission (EC) regarding the evaluation of the Common Agricultural Policy “[...] there are a very large number of various types of modelling tools available [...] But on the other hand, the applicability of these models to the questions relevant for evaluations [...] is limited” (EUROPEAN COMMISSION, 2002).

At the same time, economic modelers are concerned that their publicly available models might be over-used or even ‘abused’ (DENNIS, 2012; MCDougall, 1993), by users who lack sufficient knowledge about the model itself or impose inappropriate assumptions into their analysis. Our article focuses on post-model processing of results and results’ exploitation in the context of IA of policies affecting the agricultural sector. Two interlinked questions (see also BRITZ, 1999) seem to be relevant here: (1) how simulation results are (best) post-processed and (2) which approaches work well for analysing results.

The ease with which models can be steered and results exploited is seen as important for their more widespread use (cf. VONK et al., 2005; DíEZ and McINTOSH, 2009). There is also quite some literature presenting software to organize and visualize model results, specifically with regard to forest- (cf. FALCAO et al., 2006; MEITNER et al., 2005; SCHUCK et al., 2005) and water management (cf. BAZZANI, 2005; BANTA, 2011; WOOL et al., 2002). However, systematic comparisons with a focus on applications in agricultural economics do not seem to be available. We, therefore, examine the approaches used in three different global models, all with a long history of policy relevant applications to agricultural and food issues: GTAP (HERTEL, 1997; see also http://www.gtap.org), Aglink-Cosimo (OECD, 2007; see also http://www.agri-outlook.org) and CAPRI (BRITZ and WITZKE, 2012; see also http://www.capri-model.org). GTAP is a static Computable General Equilibrium (CGE) model, with many extensions including a recursive-dynamic one.

1 The EU has responded to the challenge by the set-up of the so-called iMAP (An integrated Modelling Platform for Agro-economic Commodity and Policy Analysis) (PéREZ DOMÍNGUEZ et al., 2008), a platform which hosts inter alia the Aglink-Cosimo, CAPRI and MAGNET models, the latter a variant of GTAP, i.e. the three models discussed in this article.

2 Another important question in this context is the choice of the software in which the simulation model is implemented. This, however, has been well-explored in the literature (BRITZ and KALLRATH, 2012). Most partial equilibrium or farm programming models applied in the European arena are based on GAMS (General Algebraic Modelling System) (BROOKE et al., 1988) and Computable General Equilibrium either relying on GEMPACK (General Equilibrium Modelling Package) (HARRISON and PEARSON, 1996) or again on GAMS. Therefore, we will not further discuss the implications of software choice in this article.
while Aglink-Cosimo and CAPRI are both agriculture-focused partial equilibrium models. Aglink-Cosimo is a recursive-dynamic, non-spatial Multi-Commodity model while CAPRI is comparative-static and links a layer of supply side models, regionalized single-country CGE and a spatial Multi-Commodity model based on sequential iteration. Each model provides a rather distinct approach with regard to post-model processing and result exploitation.

The article is organized as follows. First, we motivate that a closer look to alternative technical approaches to result analysis is useful. Next, we discuss two important approaches in post model analysis: indicators and result decomposition. Then, we review user roles with regard to applications of economic models in IA to find on which type of user to focus in our analysis. From there, we develop the main functionalities which should be covered by GUIs for economic simulation models applied for IA. Next, specifically the solutions used in the three chosen modelling platforms are reviewed as they differ considerably in their implementations and thus provide a good picture of the state-of-art, while also touching upon some other examples. Finally, we draw conclusions and summarize our findings.

2 Background

Simulations with large-scale economic models generate a large amount of numerical data, based on their differentiation by sectors/products, in space, by policy instruments and probably time. Both the size of result sets and the complexity of analysis increase further if sensitivity analysis or stochastic draws are applied, as done e.g.

with Aglink-Cosimo (BURRELL and NII NAATE, 2013; OECD, 2012) and GTAP (VALENZUELA et al., 2007). Besides computational issues related to managing large-scale data sets, communication of model results asks for a concise and coherent story to derive policy conclusions: which results are worth reporting? How can they be explained by cause-effect relations included in the model? Which regions/sectors/institutions are winners or losers and why? Equally, analysts need to remove errors and make sure that statistical outliers do not affect the results. While results generated by different models and tools will clearly differ in coverage and detail, broad strategies used in result analysis will probably not. A systematic comparison might be helpful to improve existing tools or even to arrive at common solutions as certain models might have faced specific challenges earlier than others (such as a high dimensionality in space) or had different ideas to solve some of the issues.

Modellers and analysts exploiting results have always developed and used post model visualization tools (see Figure 1). Typically, model variables are first post-processed (aggregated or presented as indicators) and smaller subsets selected. Next, these subsets are structured and customized to the user. These steps require a fair amount of a priori knowledge to decide about what aspects to include in the analysis and how to best present them with the software tools at hand. Consequently, exploitation tools of models offer such “views” often already pre-customized such that a user can simply chose from existing ones (e.g. a market balance table or a map showing changes in agricultural income), thus benefitting from past experience of other users.

3 Post Model Analysis

3.1 Indicators

Indicators can be defined as generally agreed upon proxies to measure specific impacts, condensing complex and often large sets of information. For instance, the Equivalent Variation expresses in one number “changes in well-being of the consumer” arising from price changes of products and primary factors, which themselves depend on a whole range of simulated outcomes. Similarly, total GHG (greenhouse gas)
emissions in CO₂ equivalents summarize the impact of many, interlinked bio-physical processes. Indicators are thus of key importance in communicating results, allowing users to concentrate on a few central findings in well-known definitions. VAN ITTERSUM et al. (2008) claim that “[…] scenarios are assessed through a set of indicators that capture the key economic, environmental, social and institutional issues of the questions at stake […] The indicators in turn are assessed using outputs from quantitative model components, typologies and databases”. Indicators also support cross-model comparisons, and when repeatedly used, decrease communication costs. International organizations such as the United Nations have developed indicator sets (e.g. UNITED NATIONS, 2007) while policy impact assessments often rely on indicator based approaches such as the “Common Monitoring and Evaluation Framework” to evaluate rural development programs in the EU (annex VIII of Commission Regulation 1974/2006, EC 2006). Policy-relevant models such as the ones discussed in here therefore increasingly provide indicators from different domains derived from the simulated results.

Indicators are not necessarily a model variable. Post-model calculation of indicators supports a modular structure in model development and maintenance: the structure of the simulation model can be changed without needing to update the indicator calculators, as long as their arguments (i.e. variables and coefficients) remain available. Additionally, post-model calculation allows for establishing functional relationships which might be hard to implement into the equations of the model such as, e.g., mapping an indicator into low, middle and high classes. Proper documentation of indicators is even more important than that of variables and equations, as indicators are generally designed to inform a non-technical audience.

**Policy indicators** measure the impact of policy instruments on markets. For instance, whereas most CGE models use predominantly ad-valorem import tariffs, partial equilibrium models instead often include trade restrictive measures in absolute terms and calculate ad-valorem equivalents ex-post as an indicator. Further examples of policy indicators are the boxes used by the WTO to classify domestic support (ORDEN et al., 2011) or the producer and consumer support estimates (PSE/CSE) of the OECD (OECD, 2010).

Agricultural policy design increasingly reflects impacts on the environment. Consequently, economic models try to quantify environmental impacts, typically based on **environmental pressure indicators** following the Driving forces-Pressure-State-Impact-Response model (KRISTENSEN, 2004). For instance, CAPRI calculates *inter alia* detailed nitrogen budgets (LEIP et al., 2011), phosphorus and potassium (PK) balances, GHG emissions (cf. LEIP et al., 2010) or a Life Cycle Assessment of Energy Use in agriculture (KRANZLEIN, 2008). AGLINK-COSIMO is currently working on an accounting module for GHG emissions and land use accounting layer. GTAP has invested in satellite applications dealing with environmental aspects such as GHG Emissions (BURNIAUX and TRUONG, 2002; MCDougALL and GOLUB, 2007), land use (LEE et al., 2005; LEE et al., 2008) and water (CALZADILLA et al., 2011).

A challenge provides the scale dependency of environmental indicators. Global externalities, such as GHG emissions, can be expressed in a single number (i.e. change in carbon dioxide equivalents), since the source and location of the emitter has no significance for the environmental harm. Most other externalities (e.g. ammonia emissions or leaching of nitrogen to groundwater bodies) are of a regional or local nature and can only be properly assessed in their specific spatial context. Furthermore, relationships between changes in variables simulated by an economic model and indicators are not necessarily linear such that the spatial and temporal resolution of driving results can significantly impact aggregated outcomes.

### 3.2 Decomposition and Dual Analysis

Decomposition of behavioural equations provides a powerful device to systematic result analysis. This is accomplished by relating the simulated change in a variable to changes in its determinants. Recursively over equations, one might trace all changes back to the original shock. Decomposed effects can also be aggregated to show effects on indicators, such as relating a welfare change to gains in allocative efficiency or trade effects in GTAP by HUFF and HERTEL (2000) or to decompose changes in GHG emissions in GTAP by MCDougALL and NARAYANAN (2012). GTAP (as generally GEMPACK based models) benefits here from the fact that most equations are written in percentage changes which provides already a kind of decomposition. In AGLINK-Cosimo, no formal decomposition of results is offered, but most endogenous variables are represented by log-linear functions, what allows for a straightforward interpretation of model changes through the underlying elasticities.

Models written in quantity levels using different functional forms need additional code to provide de-
compositions, such as in CAPRI where the contribution of changes in own and blocks of cross prices to the changes in production and the different demand positions is calculated (BRITZ, 2011) and aggregated from single product to group of products, and single region to a block of regions. For instance, CAPRI offers a dual analysis of the constrained optimization supply models based on Kuhn-Tucker first order conditions.

4 Overview on Selected GUIs

All three model systems analysed in-depth in here follow a strict separation of model code, data storage and graphical user interfaces. They are all global in scope, offering quite some detail with regard to, for example, products and policy instruments. However, they differ substantially in terms of nature, software used and regional resolution, as mentioned in section 1.

A first interesting observation is that the GUIs of the three systems had been basically developed by model developers and users, typically evolving over years. That might explain partly their success (see also section 4.4 on further examples): functionalities often reflect user demand, while the developer has a very clear picture about the use cases to reflect. Still, the three GUIs apply three rather distinct technical and design approaches to post-model processing (see Table 1): a more IDE (Integrated Development Environment) based concept in runGTAP, a concept more related to reporting services as found in business intelligence (CAPRI), and a kind of compromise solution realized in office software (Aglink-Cosimo).

4.1 Modelling Package Add-ons

One clear distinction between the approaches is how close the presentation of the results in tables, graphs and maps is to the native model code. Packages such as the GEMPACK graphical interface runGTAP or the GAMS IDE are add-ons to high level modelling languages, GEMPACK with its focus on General Equilibrium Models being more specialized than GAMS. Given the rather general nature of these packages, the software knows only the technical, not the logical structure of the data it handles. The exploitation facilities of GAMS or GEMPACK thus visualize symbols used in the model code such as parameters or variables; they consequently target mostly the code developer. They can hence be understood as an IDE which is also (partially) designed to support the application of the modelling package to simulations. The logical structuring of the “views” hence follows the technical structures chosen by the coder. The same can be found in variable/equation listings in the GAMS IDE or the embedded GDX Viewer. The use of these viewers hence requires familiarity with the technical structure of the model as well as with the model mnemonics.

4.2 CAPRI Exploitation Tools

The exploitation part of the CAPRI GUI (BRITZ, 2010b) chooses a quite different approach. It aims at a relatively easy to use tool to present high-dimensional data, detaching the visualization from the underlying technical data generation process, i.e. the simulation model(s) and post-model processing. All data are loaded into one multi-dimensional cube, deliberately breaking the relation to the underlying code structure (variables, equations, parameters), such that also results from different models/modules and runs might be visualized together. This principle is applied to the analysis of scenario runs, but also for inspecting the data base. It thus provides a reporting tool similar to more powerful ones available for business data bases (cf. SINGH et al., 2013), thought as a complement to e.g. the GAMS IDE needed in the development/de-bugging process.

Table 1. Overview on software solution of the three models analysed

<table>
<thead>
<tr>
<th>Model solver</th>
<th>Data handling and storage</th>
<th>GUI</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTAP GEMPACK (<a href="http://www.copsmodels.com/gempack.htm">http://www.copsmodels.com/gempack.htm</a>)</td>
<td>GEMPACK (HAR)</td>
<td>runGTAP linked to GEMPACK</td>
</tr>
<tr>
<td>Aglink-Cosimo PC-TROLL (<a href="http://www.intex.com/troll/">http://www.intex.com/troll/</a>)</td>
<td>EXCEL GAMS (GDX)</td>
<td>EXCEL in conjunction with Visual Basic</td>
</tr>
<tr>
<td>CAPRI GAMS (<a href="http://www.gams.com">http://www.gams.com</a>)</td>
<td>GAMS (GDX)</td>
<td>GAMS Graphical Interface Generator (GGIG) (Java Based)</td>
</tr>
</tbody>
</table>

Source: the authors
The data cube loaded might contain several million of non-zero entries, such that visualization in digestible bits is needed, replacing the original coding structure. That new structure is called a view and it combines: (a) filters (i.e. regions, products/sector, items, years to present to the user), (b) the pivot information (i.e. swapping possibilities between dimensions in rows, columns and additional dimensions), (c) the presentation format (i.e. table, graph, map) and (d) further properties (adding of long-texts, units and explanatory tooltips, on-the-fly data transformation, pre-defined aggregation options and hyperlinks to other tables of relevance). A good view is user-oriented by presenting a small block of related data such as elements of a market in an appropriate format (e.g. as a bar chart).

The CAPRI exploitation tools are integrated in the CAPRI GUI which offers additionally functionalities such as starting and tracking model runs, viewing meta-data, updates of data and code via Subversion software (http://subversion.trigris.org) or generating a HTML based documentation from GAMS code. These functionalities are not further discussed in the following, but show some similarities to runGTAP.

4.3 Aglink-Cosimo Exploitation Tools

The OECD-FAO model Aglink-Cosimo uses the PC-TROLL interface to run the different instances of the model chain. However, additional proprietary GUIs both support model development and result analysis. These GUIs reflect that the main use of Aglink-Cosimo which is to produce a medium-term (i.e. +10 years) baseline on commodity markets, what first requires the compilation of expert information in the form of country questionnaires for the Aglink part, second the construction of single country models with the data received and third the merge with the Cosimo model covering the rest of the world.

There are currently two GUIs in use: (1) An EXCEL based interface developed by the FAO, using a Visual Basic (VB) layer which via different TROLL drivers communicates directly with the Aglink-Cosimo database. This feature is of interest during the baseline process, for model calibration, to check market clearing conditions or to perform corrections to exogenous variables (e.g. remove outliers). (2) A more refined EXCEL tool again relying on VB tailored to model application called Aglink Model Interface (AMI), developed by DEFRA which offers beyond the simple visualization of results: (a) tracing of model variables and equations, i.e. the user can filter out model equations or search for equations that contain certain variables; (b) analysis of results, i.e. the user can have a look at yearly data and/or market balances for a certain commodity in a certain country, and (c) generation of simple counterfactual scenarios and analysis of results versus the baseline.

As Aglink-Cosimo is a recursive-dynamic model, the most natural way to look at data is time series, either in tabular aspect or graphs. But maps can be produced with the Aglink-Cosimo GUI as well. Besides the tools targeting the model developer and analysts, it provides to official users a ready-to-use exploitation tool to access information through the Aglink-Cosimo collaboration site (http://www.agri-outlook.org).

The GUIs are specifically designed for Aglink-Cosimo and cannot be extrapolated to other model systems. Changes to them require a fairly good knowledge of Visual Basic, creating a certain dependency on IT specialists.

4.4 Further Examples of GUIs

In the last years, at least two large-scale EU research projects SENSOR (Sustainability Impact Assessment: Tools for Environmental, Social Multifunctional Land Use in European Regions; HELMING et al., 2008; see also http://www.sensor-ip.org) and SEAMLESS (System for Environmental and Agricultural Modelling: Linking European Science and Society; VAN ITTERSUM et al., 2008; see also http://www.seamless-ip.org) aimed to develop IA tools in the wider context of agricultural policies. SEAMLESS developed the SEAMLESS-IF, an Open-MI based environment to link models (MOORE and TINDALL, 2005) which comprises a web based GUI which also support different stages of IA processes (JANSSEN et al., 2009). SENSOR developed the so-called SIAT tool (Sustainability Impact Assessment Tool, VERWEIJ et al., 2012). Both projects devoted considerable resources to develop GUIs. The only application of SEAMLESS-IF known to us provides a test case during the project’s lifetime (BEZLEPKINA et al., 2010); it seems not to be used afterwards for applications of the components developed, improved or linked in SEAMLESS. There is further anecdotic evidence of abandoned links between economic simulation models and GUIs. A Swiss team had a rather powerful GUI developed for their agricultural sector model SILAS (MALTIUS et

---

3 Currently 14 countries are part of Aglink: USA, EU, Canada, Australia, New Zealand, Japan, South Korea, China, Russia, Mexico, Argentina, Brazil, Switzerland and Norway.
al., 2000) which was later given up. The generic GUI tool GSE (Dol., 2006) still in use, updated and enhanced by LEI and used with several economic simulation models was added to the CAPSIM model (Witzke and Zintl, 2005), but afterwards not applied. In all the cases mentioned, the simulation model (or components thereof) continued to be used, but no longer in conjunction with the GUI.

Based on general reasoning and these examples, we propose three conclusions. Firstly, not only the original development, but also maintenance costs of GUIs matter. Many economic simulation models are continuously developed over time, e.g. by adding new modules or increasing their resolution or scope with regards to products, regions or policy instruments covered. These changes must be typically reflected in the GUIs, and if they cannot be easily implemented by the model developers themselves, external expertise must be hired, causing costs (contracting, communication with IT experts etc.) and often delays for new model releases. Secondly, GUIs should not considerably drive up computing time during model development and application. And thirdly, GUI should not (drastically) restrict the user in using the model code. Not all GUIs discussed strictly separate the GUI code from that of the economic simulation model. In the extreme case, model runs are not possible without using the interface.

5 Result Exploitation in Selected GUIs

Large-scale models produce a fair amount of results that need to be analysed frequently by few people. On the one side, model developers need to continuously look at results during the development and debugging phases of a model project. On the other side, policy and market analysts working with model results need to see the information condensed in few numbers that can quickly be understood and made available in the form of a policy briefing or fact sheet. For this, having the appropriate graphical tools of exploitation at hand is very important, so that data can be quickly accessed and model results processed to be presented to the respective clients and user groups.

The typical application of a GUI when assessing scenario output consists typically of three probably overlapping steps. Firstly, the analyst will check if the intended scenario was implemented. Here, policy indicators are helpful, e.g. reporting budgetary outlays, ad-valorem tariff equivalents, or policy instruments classified according to the WTO boxes or the OECD’s PSE concept, especially in Partial Equilibrium (PE) models with a detailed representation of agricultural and food policy instruments. Next, as a further part of quality assurance, major impacts with regard to direction and size are analysed, comparing against a priori expectations based on economic knowledge and market intelligence. Here, overview tables with results aggregated across regions and products/sectors are useful, clearly in combination with relative changes against the baseline. Once these checks are successfully passed, the analyst will typically start to look for results that are worth to be reported, which are often larger relative changes for important products and regions. That step typically also involves a search for possible outliers. Here, visualization of distributions across products or space via histograms or statistical outlier detection are helpful, they are often combined with the use of maps where the human eye can easily detect exceptional pattern. Finally, selected results are either reported directly within the text or as separate tables, graphs or maps. In most cases, tables and graphs are manually formatted and comprise carefully selected results, and are not directly taken from the GUIs.

5.1 Filtering Capabilities

Filtering reduces the information to be inspected by the user, to only view those results which matter most for the topic analysed. Defining an appropriate filter requires ex-ante knowledge or a data driven approach to choose the most important data for the specific question to analyse. Typically, larger parts out of a result set produced by a model can be masked out by rather trivial rules. But there might remain a larger grey zone with data where it is not clear if they matter or not for a specific question. In that case, predefined filters which allow expanding or shrinking the analysed data set might be helpful in presentation of results.

Analysts frequently develop their own filters when working with results and repeatedly apply them. Examples for such “filters” are tables used in project reports or spreadsheets developed for analytical purposes. Combing the variables from model (solutions) which provide a market balance sheet is hence also a “filter” in our understanding. In many cases, such reports comprise algebraic transformations of the original data, such as calculation of shares or of absolute or relative differences. These user developed filters represent structural knowledge applied successfully in
result analysis. Learning from them and “storing” them for later re-use can guide other users in result analysis, safe time, prevent information overload and avoid searching for or seeing spurious relations not embedded in the equation structure of a tool. A tactic followed by CAPRI and certainly other modelling tools consists in the development of tools which allow storing such filter for later re-use.

The main approach in CAPRI to support result analysis is the application of pre-structured reports, as clearly found in many other models and tools. In CAPRI these reports are grouped to themes, where each theme comprises a list of reports. As an example, the reports under the theme “Markets” comprise different types of market balances and other views related to the global trade model of CAPRI. Besides showing only a few pre-selected elements from the whole result data set, each report is linked to an appropriate view type (i.e. table, map or chart).

The basic approach of masking out data by filters is naturally also found in other viewers. In the case of GTAP the user is required to also understand the individual symbols used in the model, since no long texts are allowed by GEMPACK.

Statistical filtering removes data points during viewing based on data driven approaches, typically by hiding data points from the view below certain numerical thresholds. In CAPRI, the GUI allows to show basic statistics in any view (e.g. mean, median, variance, and quantiles) and to use different statistical outlier methods to hide any “normal observations” and show only few “suspicous” ones. Concentrating on statistical outliers, sorting by relative changes or hiding any results with relative / absolute changes below chosen thresholds provide additional possibilities to the researcher to detect bugs or present useful information. Equally, a histogram view can be added to tables, graphs or maps to visualize the distribution.

The filtering of information in Aglink-Cosimo is not an easy task, mainly due to the underlying software. PC-TROLL is mainly statistical software which also includes a solver (i.e. mathematical algorithm) for large systems of simultaneous equations such as Aglink-Cosimo (around 23,000 equations) in a recursive-dynamic fashion. PC-TROLL works only with two dimensions (series, observations), not offering the possibilities to break up the series dimension additionally, e.g. into regions, product and items in a pre-compilation phase. This is radically different in GAMS or GEMPACK, which provide an indexed language right from the beginning. Whereas in GAMS and GEMPACK equations can be presented in a template fashion which is vectored on demand, PC-TROLL requires writing each equation out individually. The user can naturally improve readability using strict mnemonics (e.g. MEX_SU_EX describing the series for the region Mexico, the product sugar and the item exports), but PC-TROLL remains unaware of that structure and cannot exploit it. As a consequence, mapping routines have been developed for Aglink-Cosimo to convert the original two dimensional cube (variable - year) into a four-dimensional representation (region - product - item - year). The Aglink-Cosimo viewers then allow filtering of data/results based on the resulting multidimensional cube.

5.2 Condensing and Normalizing Information

Aggregation is the simplest way to condense information. In CAPRI, aggregation is rather systematically done for almost all results: (a) over the regional scales in the model (i.e. farm type => sub-national region => nation => global region => globe); (b) over single outputs / inputs / activities (i.e. market balance results, prices, etc.); and (c) over items (i.e. aggregation of different border protection measures to ad-valorem equivalents or demand components to total demand). The aggregation rules are typically implemented in GAMS code to apply appropriate weights, which is eased by set driven concept of GAMS allows where the rules can be coded independently from the actual list of components and aggregates. Similar to CAPRI, post-model aggregation in Aglink-Cosimo is done automatically based on fix aggregation routines pre-programmed in PC-TROLL, over regional aggregates and groups of commodities. These aggregation routines reflect the needs of users and/or match the pre-defined formats required by the official OECD publication and statistical departments.

The approach in GTAP differs from that in CAPRI and Aglink-Cosimo, since aggregations are typically performed pre-model. A flexible aggregation over regions and sectors is only feasible in GTAP, where data, parameters and model equations follow a template format. Pre-model aggregation reduces solution time and expected errors in input data, as long these errors are not highly correlated. It also lowers the number of small absolute values including cost and trade shares which helps model solution. However, pre-model aggregation of policy instruments such as production or tariff rate quotas, administrative or minimum prices, as well as subsidies subject to ceilings in
value or quantity terms, require approximating them as ad-valorem equivalents. The latter is the major reason why even fully template models such as CAPRI abstain from flexible aggregation. In a not fully template system such as Aglink-Cosimo, where behavioural, price transmission or market clearing equations might differ between regions and markets, flexible aggregation is not possible.

In CGE analysis such as with GTAP, regional/sector/commodity aggregation might hence change from application to application. When the user has already defined an appropriate (dis)aggregation level for analysis in the actual model which reduces the dimensionality of the results, post-model aggregation might not be necessary. Maintaining an unchanged regional and commodity list with related aggregate definitions such as in CAPRI or Aglink-Cosimo has, however, also its advantages: experiments run by different teams/projects and relating to different shocks can be compared to each other, often even across different model versions. Post-model aggregation by commodity, activity and/or region in CAPRI and Aglink-Cosimo provide hence an alternative to flexible aggregation. To give an example, the regional aggregation in the CAPRI trade model comprises two aggregation layers, one by geographic aggregates (e.g. Africa, South America, etc.) and one from a trade policy perspective (e.g. LDCs, developed countries, etc.). Flexible pre-model aggregation will also introduce new mnemonics for each new aggregation, while results obtained will depend to some extent on the chosen aggregation.

Normalization of results, e.g. to calculate shares or to derive relative or absolute changes against a comparison point, is an interesting option to condense information. That is the usual way to present results in GTAP and many CGEs were prices represent indices. In CAPRI, the user can add relative or absolute changes against any item in any dimension to a view; that can also easily achieved in the MS-Excel-based interfaces of Aglink-Cosimo.

5.3 Graphical Presentation of Results

All three systems offer graphical exploitation possibilities. Whereas the CAPRI GUI includes a wider set of graphs (line, point, bar, pie, spider, histogram, deviation renderer, box & whisker), runGTAP features a viewer called “charter” for line and bar charts. Analysts working with CAPRI and GTAP tend to produce graphs for reports with office software to overcome the somewhat limited formatting possibilities in these GUIs. Instead, the Aglink-Cosimo interface realized in VB, intensively used by analysts to inspect data and results during the baseline generator process, benefits from the high flexibility to format graphs in MS-Excel such that the views shown in the GUI can also be directly integrated in reports.

Maps are also a powerful exploitation tool. In the case of Aglink-Cosimo, no maps have been used up to date due to territorial disputes, although mapping software is available. GTAP comprises a basic tool for maps (ShadeMap, http://www.monash.edu.au/policy/shademap.htm). With now ~300 regional units for Europe or even ~200,000 1x1 km clusters in the spatial disaggregation tools (Leip et al., 2008) in CAPRI, maps are widely used for analysis in reports. As most CAPRI users are not used to GIS mapping software, the CAPRI GUI tries to offer the necessary formatting and classification possibilities to produce camera-ready maps for publications. These maps can be accompanied by coloured histograms which report basic statistics and visualize the chosen classification over the mapped data set (such as quintiles, equal interval, nested means, and natural breaks).

5.4 Experiences from Model Applications

By analysing journal publications and reports published in the last few years, we conclude that model results are most often presented in tables, and frequently at more aggregate level (product and country/regions). The PE models CAPRI and Aglink-Cosimo typically report both absolute quantity and price levels, as well as relative differences, while GTAP typically reports relative changes, and more often also monetary values (= quantities times price). Only CAPRI uses maps regularly, these being typically generated directly with the GUI (e.g. changes in land use and greenhouse gas emissions). Graphs are seldom used and more often found in CGE analysis where often rather aggregate relative changes, e.g. across global regions or sectors, are visualized.

From the various training sessions organized by the CAPRI network, it emerged that calculation of relative changes against the baseline is perhaps the most important feature when results are quickly scanned; the same holds for pre-defined aggregations over products and regions. In interactive scanning of results, maps are more often used than in published analysis, as they can be produced by the CAPRI GUI rather quickly and allow a very fast assessment of spatial pattern of changes, including possible outliers. The possibility to drill-down into more detailed results (e.g. decompose an income change in an activity into changes of revenues and of different cost positions) is...
also quite useful when interpreting results in reports. The more formal decomposition approaches, e.g. into own and cross-price effects, are still not widely used by CAPRI analysts. Graphics are mostly used in connection with the construction of the baseline where projected results for different time points and a range of data providers are compared against the ex-post development.

In the GTAP framework, several applications allowing for further analysis of results are available; for example, in addition to the income decomposition mentioned above, all the model equations with multiple terms can be analysed in this way. Trade changes may be decomposed into domestic penetration or import expansion effect in the destination country and substitution effect entailing trade diversion from and to other countries. Welfare decomposition is yet another application. Policy-makers are often puzzled when the researchers talk about broad welfare change measures, which are typically based on Equivalent Variations; this has been broken down in the GTAP model into multiple components, such as allocative efficiency, terms of trade, etc., which can be further decomposed to input, sectoral and regional level. Graphics may also be easily created within the GTAP solution and data files, but maps are difficult to construct. This is an area of work for the future for the global CGE models.

In Aglink-Cosimo applications, results are typically presented as tables and graphs. Since one of the main applications of this model is the production of medium-term agricultural market projections by different governments and international organisations, applications frequently concentrate on the development of agricultural market balances (i.e. supply, demand and trade) over time. As a consequence, applications concentrate on how (and when) policies are likely to affect a specific market. Aggregation of results in Aglink-Cosimo is an issue, since it is typically included as part of the model and, therefore, lacks the flexibility of more modular post-model processing approaches. In this field, the potential for using more modular approaches through the use of flexible GUIs is large.

6 Summary and Conclusions for GUI Design

The GUIs of GTAP, CAPRI and Aglink-Cosimo, economic simulation models regularly used for policy relevant applications in the context of agriculture and food, provide examples for three rather different basic approaches. GTAP builds on GEMPACK tools which are linked to a specific modelling software package. The data viewer, with its tools, has more the character of an IDE, and is strictly linked to the technical presentation of symbols in the model. CAPRI has opted to develop exploitation tools which are much closer to reporting tools linked to a Data Base Management System. These tools are technically and formally detached from the data generation process, i.e. the economic model and its post-processing steps. Aglink-Cosimo finally shows exploitation tools formally separated from the underlying software but specifically developed to serve this model. The three GUIs offer the basic visualization possibilities (i.e. tables, graphs, maps), however, with quite some differences in detail. Aglink-Cosimo uses MS-Excel as a carrier for tables and graphs, and therefore benefits from its versatile formatting options. Its views are also integrated in certain work steps of the model such as baseline generation with which they can easily interact. Similar to GEMPACK, the Aglink-Cosimo GUIs can show the equation which defines a certain variable currently viewed.

The CAPRI exploitation tools put a focus on content-guided access to results by grouped thematic views, on interactive navigation, e.g. by hyperlinks between logically connected tables, on adding meta data (i.e. measurement units, explanatory texts, date of release, user type, etc.) and have integrated data driven approaches such as basic statistics, outlier detection methods and recently a Machine Learning package (WITTEN et al., 2011). CAPRI also features the most versatile mapping possibilities. However, a link to the underlying technical structures (i.e. equations, variables) is not integrated in the viewer, but only available via a separate utility.

This article does not aim to propose improvements for the individual viewers, but rather raises the question if the best of the different worlds cannot be combined into one shared GUI tool. Interestingly, two generic, license-free tools for GAMS based models exist: firstly the so-called GAMS Simulation Environment (GSE) (DOL., 2006), where the GUI is set-up by introducing specific “tags” as comments into the GAMS code, and secondly GGIG (BRITZ, 2010a), developed from the former proprietary CAPRI GUI, which combines the functionalities of the CAPRI exploitation tools (e.g. viewing of GDX files or generation of maps, etc.) with the possibility of building a GUI to steer economic models by the user just by editing a text file in XML format. Both tools are al-
ready successfully used in a number of projects (e.g. GGGIG for a single farm model, LENGERS and BRITZ, 2012; a river basin model, KUHN et al., 2014; recently the Aglink-Cosimo model; and a GAMS variant of GTAP, BRITZ, 2014). GSE and GGGIG thus underline that generic solutions are feasible. Both allow model developers to update the GUI to their needs (e.g. to reflect changes in the product or regional lists or to add new views).

These generic approaches avoid costly model specific GUI coding and maintenance efforts and decrease learning costs for researchers switching models. Joint maintenance and further development of a generic solution would also reduce the obvious risks related to the fact that currently most model (family) related exploitation tools are only programmed by one or at best a handful of people. It is however doubtful that each cherished functionality e.g. now found in the proprietary GUIs of GTAP and Aglink-Cosimo could be integrated in a generic one. And clearly, a software change such as moving to a different GUI devalues past investments in human capital, which might frustrate researchers and other staff who have spent considerable time to familiarize themselves with a specific software solution and now have to learn another one, and provokes immediate costs (training, GUI set-up, licensing costs or contribution to costs of shared development), with perhaps unclear future returns. But at least for newly to develop tools, a close look at the existing generic packages might pay off.

An intermediate solution, similar to what is found for Aglink-Cosimo, would consist in developing interfaces to the existing viewers, so that different exploitation tools can be used with the simulation results of a model, giving model users more freedom and flexibility.

References


The former proprietary GUI of CAPRI was already moved to GGGIG, and almost all functionalities had been ported.


Contact author:

DR. WOLFGANG BRITZ
Institut for Food and Resource Economics, University Bonn
Nüßallee 21, 53115 Bonn
e-mail: wolfgang.britz@ilr.uni-bonn.de

All rights reserved www.gjae-online.de