Grant Agreement no. 308371
ENV.2012.6.3-2 - Policy Options for a Resource-Efficient Economy

- Collaborative project -

D3.4
Report on the linking of GINFORS and LPJmL

WP 3 - Scenarios and modeling of policy implementation for resource efficiency

Due date of deliverable: Month 18

Submission date: 10/03/2014

Start date of project: 1st October 2012 Duration: 42 months

Lead beneficiary for this deliverable: GWS

Last editor: Bernd Meyer

Contributors: Tim Beringer, Martin Distelkamp, Frank Hohmann, Bernd Meyer

This project has received funding from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement No 308371.

Dissemination Level		
PU	PAB members	X

History table

Version	Date	Released by	Comments
1	10/03/2014	Bernd Meyer	Circulated to POLFREE Policy Advisory Board for comment
Version	Date	Released by	Comments by Liz Goodwin and Jill Jaeger
2	31/03/2014	Bernd Meyer	

Table of contents

1. Introduction

The modelling of resource use is discussed here in the context of an adequate integration of the demand for resources in physical terms into the economic system, which among other things means that the interaction of supply and demand has to be considered. In the case of fossil fuels, metals and non-metallic minerals, the supply side is to a large extent part of the economic system. Extraction costs and investment needs are known as economic variables, only the physical availability of the resource is determined outside the economy, but it has the advantage of being given exogenously. In a sufficiently detailed economic modelling approach, it is possible to depict the interactions of supply and demand to give a full picture of resource use. In most cases the supply side is represented by the exogenous price of the resource, the rise of which signals the growing scarcity of the resource in question. Some models have demand and supply curves for fossil fuels and are able to calculate the resource prices endogenously.

The integration of agricultural land use for different types of biomass and the use of water in economic development is much more complicated, because the physical availability of the resources is determined to a large extent outside the economy and is not exogenously given. The supply of water depends on the geographical unit's hydro-geological and bio-geochemical characteristics, which change over time and which are interrelated with the economic system, which itself also creates the demand for water. "The interaction between the hydrological and the economic realm works both ways: water is transformed for economic use and the impact of economic use on water availability and quality consequently has implications in both the short and long term for the transformation process to modify water for economic use" (Brouwer & Hofkes, 2008, p.17). A similar situation arises for agricultural land use: Economic decisions determine which crops are produced where, but the soil quality, water availability and other dynamic physical factors determine the growth of crops and the result of the economic activities. Furthermore, soil quality and water availability are influenced by economic factors.

Braat & van Lierop, 1987, differentiate between the holistic and the modular approaches to integrated hydrological economic modelling. The holistic approach solves one integrated model with hydrological and economic equations simultaneously (Ward & Pulido-Velázquez, 2008, Cai et al., 2008, Pulido-Velazquez et al., 2008) whereas in the modular case the hydrological and the economic model exchange data, which are for the other model exogenous (Volk et al., 2008, Jonkman et al., 2008, Barton et al., 2008). Brouwer & Hofkes, 2008, add the CGE modelling as a third approach. This seems to be problematic. Firstly the term "CGE" stands for a sectorally disaggregated macroeconomic model based on strict neoclassical assumptions, and which is parameterised by calibration for the data of one observation. The authors stress the total and complex modelling approach, which includes the sectoral interdependency in comparison to those rather simple economic systems that have been used by the existing “holistic” and “modular” approaches. Sectorally disaggregated total models of the macro economy can also be based on the assumption of the bounded rationality of the agents and have econometrically estimated parameters like the GINFORS model. But it is not only the name which seems to be problematic for the third category. Also such big and complex models as the CGE models or models like GINFORS could be part of a “holistic” or a “modular” approach. Furthermore, for adequate modelling of water and agricultural land use, such complex economic models are necessary, and a two-way link with a bio-geochemical model – in our case it is the model LPJmL - is essential, which means that either a “holistic” or a “modular” approach has to be chosen.

The simultaneous solution in the “holistic” approach is, of course, the most appropriate. But we have to mention that all of the holistic models discussed in Brouwer & Hofkes, 2008 are regional models depicting the relations in a river basin. Compared to that, the bio-geochemical model LPJmL and the economic model GINFORS are huge global models with a deep country and sector structure. This does not allow for a simultaneous solution of both systems. But with the assumption of a lagged response by nature to changes in the economic system, we can reduce the complexity of the task since both models can be solved period by period. This assumption is very realistic, because it is just the period of production which occurs between the decisions about the seeding and the harvesting of crops.

In each period, the economic model GINFORS starts its calculations based on data from the bio-geochemical model LPJmL from the previous period, delivers data to LPJmL, which solves it for the same period. In a technical sense we then have a modular approach with an independent solution of both models and an exchange of exogenous data. But from the subject matter and in a long-run perspective, we have a simultaneous solution, since only the appropriate assumption of the production period gives the recursive solution structure.

The literature discusses the following problems which arise from linking bio-geochemical and economic models (McKinney et al., 1999), and a short look at it shows that a hard link between LPJmL and GINFORS might be successful.

Bio-geochemical models are usually solved period by period using standard routines for the solution of linear or nonlinear systems of equations. Economic models often use optimisation techniques over a long period, which would not allow the procedure just described. The GINFORS model is not an optimisation model and therefore can be solved period by period as with LPJmL.

In hydrological and other bio-geochemical models, the geographical unit is often a river basin, whereas the economic models define the regions with administrative criteria, resulting in countries or provinces. In our case, the model LPJmL divides the surface of the earth into cells of 10*10 km; GINFORS divides the global economy into 38 countries and the region "Rest of the World". Since the cells of LPJmL can be aggregated into the countries of GINFORS, there will be no problem concerning the geographical units.

The timescales of bio-geochemical and economic models are different. GINFORS has a period of one year. LPJmL refers to months, but the data delivery to GINFORS can be aggregated into the year. Furthermore, it is plausible that the economic data delivered to LPJmL can be used for all sub-periods of the year.

Based on the description of the modelling of agricultural land use, water and biomass in deliverable D3.2, the paper at hand presents in detail the technical solution of the linkage in Chapter 2. It further analyses the interfaces of the models in Chapter 3. The most interesting aspect of the linkage with regard to the content is the reaction of GINFORS to restrictions that the solution of LPJmL may mean for the solution of GINFORS. Some expected effects will be discussed in Chapter 4. Some conclusions in Chapter 5 close the paper.

2. Technical solution

The process of linking huge models such as GINFORS and LPJmL involves both *logical* and *technical issues* which need to be resolved.

Logical issues are related to questions such as which data have to be transferred between the models, how to find the appropriate position in the model where the data transfer should take place, and also how to deal with stability problems which might be introduced by exogenous data. These questions are covered in Chapters 3 and 4.

Technical issues are manifold as well. The models were initially developed as "stand-alone" solutions with each of them having their own dataset, programming language, file formats and solution algorithm. The linkage between these models inevitably requires some changes to each of them. The technical solution has to be carefully designed to make sure that the impact on each model is as low as possible and that the stand-alone version of each model can still be easily maintained.

As already pointed out in Chapter 1, the size and complexity of both GINFORS and LPJmL do not allow for a simultaneous solution. Instead, the solution algorithm implies that for each year *t*, both models may be executed in sequence. For each *t*, model A [LPJmL] solves first and generates output data which serves as input data for the other model B [GINFORS]. Model B then tries to solve for *t* as well and – in the case of success – generates output data which will be used as input data by model A in *t+1*. This process has to be repeated until the final year of the simulation time span is reached. If a model was not able to solve in *t* for some reason, the simulation process has to be stopped before the final year has been reached.

From a technical point of view, controller software is needed which is capable of initiating, controlling and evaluating the model execution over time. Since this project not only involves the linkage between GINFORS and LPJmL but also between LPJmL and EXIOMOD, it is highly desirable that the software controller may be applied to both linkages.

The main tasks of the controller software are "injecting" the current year *t* into the execution procedure of each model GINFORS, LPJmL and EXIOMOD and evaluating the return value which indicates whether a model was able to solve successfully in year *t*.

2.1 Model parameterisation

For each model-building environment (GINFORS, LPJmL and EXIOMOD), the controller software needs to know how to pass the current year *t* and also how to evaluate the result of the execution. Although each model can be executed within the Microsoft Windows operating system, the technical solution has to carefully address the differences between the model-building environments.

GINFORS is built in the programming language C++ on top of the model-building framework *Solve* which has been developed by GWS. *Solve* provides mechanisms for evaluating command line parameters as well as taking parameters from a parameter file.

The value for year *t* could be passed to the model by providing two command line parameters *startYear* and *endYear* set to the same value *t*, e.g. 2010:

```
Model.exe startYear=2010 endYear=2010
```

Alternatively, the same information could be passed to the model by modifying the default parameter file *solve.conf*. The format of the file is a called *Windows Ini Format* which is also widely used by programs running under the Linux/Unix and Mac operating systems. Parameters are given by *key-value pairs* separated by an equal sign ("="). Related parameters may be grouped into *sections* which are indicated by brackets "[" and "]". An excerpt of the parameter file is shown below:

```
[solve]
startYear=2010
endYear=2050
maxIterations=300
model=GINFORS3
title=Baseline
```

The controller software has to set the values of *startYear* and *endYear* to the value of *t*:

C++ programs must contain a *main* function. This function must return an integer value to the operating system which indicates the result of the execution. Such programs usually return 0 (zero) if the execution has been successful and other integer values if some sort of error or malfunction has occurred. The simplified main function of GINFORS is shown below:

```
int main() {
 try {
      ginfors3 model;
      model.configure("solve.conf");
      model.run();
      cerr << "done.. " << endl;
 } catch(const exception& ex) {
      cerr << "\nERROR: " << ex.what() << endl;
      return 1;
 }
      return 0;
}
```

The controller software will have to stop the simulation process if it encounters a value for *result* which is non-zero, or otherwise continue the simulation process. A non-zero value may occur if either the system software has detected some sort of error (e.g. a missing database file) or if the model could not find a solution and exceeded the maximum number of iterations *maxIterations* per year as set in the configuration file solve.conf.

LPJmL is written in C which is the predecessor of C++. The mechanisms applied to GINFORS may be used with LPJmL as well. If the controller software encounters a non-zero return value, it will interpret this information as an error or malfunction and therefore will halt the execution of the simulation process. A value of zero indicates that the execution has been successful and control may be passed to GINFORS for the current year *t*.

LPJmL uses a parameter file which not only uses a different file format than the GINFORS parameters file but also has many more entries. The interesting part of this file is shown below:

```
#ifndef FROM_RESTART
5000  /* spinup years */
/* exclude next line in case of 0 spinup years */
30   /* cycle length during spinup (yr) */
1901 /* first year of simulation */
1901 /* last year of simulation */
```

```
NO_RESTART /* do not start from restart file */
RESTART /* create restart file: the last year of
simulation=restart-year */
restart/restart_1900_nv_stdfire.lpj /* filename of restart
file */
1900 /* write restart at year; exclude line in case of no
restart to be written */
#else
390  /* spinup years */
/* exclude next line in case of 0 spinup years */
30 /*cycle length during spinup (yr)*/
1901 /* first year of simulation */
2009 /* last year of simulation */
RESTART /* start from restart file */
restart/restart_1900_nv_stdfire.lpj /* filename of restart
file */
RESTART /* create restart file */
restart/restart_1900_crop_stdfire.lpj /* filename of restart
file */
1900 /* write restart at year; exclude line in case of no
restart to be written */
#endif
```

From a technical point of view, it is not necessary that the controller software is capable of interpreting every entry in this file. It is sufficient that the controller is able to locate and alter the two lines marked in grey.

EXIOMOD is built using the model-building software *GAMS*, which comes with a user-interface for both editing and executing a model. Alternatively, a GAMS model can be executed by passing the name of the model file as a command line parameter to the main executable *gams.exe*. Additional parameters may be passed through key-value command line parameters:

```
gams.exe modelname key1=value1 key2=value2
```

A GAMS model may also be parameterized by providing a parameter file with formats such as txt (text), csv (comma-separated value) or Excel.

Evaluating the result of a GAMS model execution proves to be different from GINFORS and LPJmL, Table 1 shows the list of possible *program return values* provided by the GAMS Compiler and Execution system (cmexRC):

Return value	Description
0	normal return
1	solver is to be called; the system should never return this number
2	there was a compilation error
3	there was an execution error
4	system limits were reached
5	there was a file error
6	there was a parameter error
7	there was a licensing error
8	there was a GAMS system error
9	GAMS could not be started
10	out of memory
11	out of disk

Table 1: Program return values provided by the GAMS Compiler and Execution system (cmexRC)

This list only provides information about possible *system* errors. It is not possible to evaluate whether the model was able to find a solution or not. A return value of 0 (zero) “normal return” only indicates that the model was executed. This value is returned even if the model was not able to find (an optimal) solution. To quote the GAMS Interfaces Manual: "*Using the return codes allows the calling program to find out if there were any compilation or execution errors or other reasons why the GAMS job could not be completed. It is noted that return codes do not say anything about a model inside the GAMS job: the model may have been infeasible or may have failed in another way while the return code says all is fine. In fact there may be multiple solves in a GAMS job, so even conceptually it is not possible to return solution status codes in the return code.*”[1] The controller software therefore may use the return code only for detecting possible technical problems.

During execution, a GAMS model outputs a comprehensive set of status information to the screen. This output can be redirected to a file for further investigation. The most important

---

[1] GAMS Interfaces Wiki, 2014

information with respect to model linkage is the *status report*. The following output comes from a simple transportation example (Rosenthal, 2013, p.23):

```
S O L V E  S U M M A R Y
MODEL      TRANSPORT      OBJECTIVE      Z
TYPE  LP        DIRECTION      MINIMIZE
SOLVER     BDMLP     FROM LINE      49
**** SOLVER STATUS   1 NORMAL COMPLETION
**** MODEL STATUS    1 OPTIMAL
**** OBJECTIVE VALUE          153.6750
RESOURCE USAGE, LIMIT     0.110     1000.000
ITERATION COUNT, LIMIT    5         1000
```

The most important line is marked grey: The *solver status* code returns the result of the model solving algorithm. Any value different from 1 should be interpreted as an error. This information could also be written to a separate file *result* by adding the following lines to the model code:

```
file result ;
put result ;
put 'ModelStatus', transport.modelstat:2:0 /;
put 'SolverStatus', transport.solvestat:2:0 /;
```

The *result* file then just contains the following lines:

```
ModelStatus 1
SolverStatus 1
```

Table 3 summarises possible values for program return codes, solver status code and actions the controller software has to take[2]:

---

[2] It might be necessary to evaluate the model status value as well if model linkage causes infeasible solutions which have to be rejected.

Return Code	Status Code	Action
0	1	Continue simulation
0	<>1	Abort simulation
<>0	N/A	Abort simulation

Table 3: Values for GAMS return codes, solver status code and actions the controller software has to take

2.2 Data exchange file format specification

The data exchange has to be negotiated between the linked models. This involves both *logical* (covered in Chapter 3) and *technical* issues.

Technical issues are related to the fact that GINFORS, LPJmL and EXIOMOD are built in different model-building environments and therefore have different database, import and export file formats. E.g. GINFORS' database file format is proprietary and cannot be read by or written to any other model-building environment.

To select an appropriate file format for data exchange which can be used with all three models without too much programming effort, the first question to answer is whether the file format should be *text* or *binary. Text file* formats are human-readable, meaning that such files can be edited by users with any text editor (large files may require professional editors, though). For bigger data sets, these files tend to be much bigger than their binary counterparts and thus take more time to process. Binary file formats cannot be read by humans but tend to be very compact and can be processed very fast even with big data sets.

For this project, a simple text file format should be preferred over a binary file format for two reasons. Firstly, at any time it must be possible to easily check and verify the data which have been passed between the models. As already pointed out, with a text file format any text editor may be used to accomplish this task. Secondly, the amount of data which needs to be transferred between the models is rather limited, so that size or processing speed plays a minor role.

The suggested data exchange file format has the following specifications: Each model should write out its data as a CSV (comma-separated values) text file. The decimal point has to be a "." (dot); values have to be separated by a semicolon. The first value (or column) on each line should be a variable name. The second value should be an integer value which indicates how many observations are available for this variable. The remaining values on this line are the observations for this variable. A file may contain many variables but only one per line. Numbers should be stored in fixed notation without separators. An example file is given below:

```
Variable1; 5; 12.124; 11.511; 13.122; 12.162; 13.151
Variable2; 2; 1452.154; 1418.881
```

The data for each year needs to be stored in a separate file. The file name contains the name of the model which produced the data and the four-digit value of the year, e.g. GINFORS2005.csv. Each model should use its own directory to store its data files.

2.3 Model linkage controller software

The controller software's responsibility is to start the linked models and to evaluate their return codes. The execution of the models does not involve any user interaction; it is therefore not necessary to provide a graphical user interface for the controller software. The controller software will be developed as a terminal program which takes its settings from a configuration file. The format of the configuration file will be the same as the one that is used by the GINFORS model (Microsoft Windows Ini file format, see Chapter 2.1).

The configuration file must contain a value for each *start* and *end year* of the simulation process. Being able to set these values independently is especially useful for debugging purposes.

For each model, the controller software needs to know the directory name and also the name of the main executable that has to be launched.

As already pointed out in in Chapter 2.1, each model has its own configuration file format. From the controller software's point of view, it does not need to be able to interpret the content of these files except for those lines which contain the start and end year values. The values for the start and end year have to be the same for each iteration of the simulation process, because each model computes only one year before passing control back to the controller software. For each linked model, the controller software will read a template configuration file at the program start which contains a special token @@y. During the simulation process, this token will be replaced by the current year before the model executable is launched.

Every time the controller software launches one of the linked models, it has to evaluate whether the model has succeeded in finding a solution or not. For the GINFORS and LPJmL models, the controller has to check the program return code. A value of zero indicates success; any other value indicates an error. For EXIOMOD, the result file needs to be evaluated by locating the line which contains the solver status. Only if the result file contains "SolverStatus 1" was the model able to solve successfully.

3. Model-interaction with regard to contents

To handle the technical challenges of a hard link between two very capacious and detailed global forecasting and simulation models is one point. But besides these technical challenges, one also has to reason about the “right” model-interaction with regard to contents. Concerning this matter, the preceding Deliverable D3.2 already discussed a first version of the new structures that are needed within the economic model GINFORS to analyse the interlinkages between global socio-economic developments, crop production and water and land availabilities. Whilst these descriptions in D3.2 were mainly based on research on data availabilities and the abilities and needs of the economic model GINFORS, the following specifications show a more integrated view from both modelling sides. Therefore in a first sub-chapter, the content of the data exchange between the models is specified. Chapter 3.2 gives a short description of the bio-physical model LPJmL with special emphasis on the implications of information (data) that comes from the GINFORS model. Since the concept presented here contains some improvements compared with the modelling concept of D 3.2, we have described these amendments in sub-chapter 3.3..

3.1 Content of the data exchange between the models

Table 5 depicts the content of the data exchange between the two models. After each model solution of LPJmL for one year, 546 values (for production quantities and for the total irrigated area) are transferred to GINFORS. Based on this input, the iterative solution procedure for the same year within GINFORS starts. At the end of this process, 702 values (for agricultural land use, water abstraction and land scarcity) are transferred to LPJmL, and the combined model solution for the next year can start.

Information flow	Direction
Production quantities (in tonnes) for 13 crop types in 38 countries and Rest of the World for the actual year	LPJmL ⇨ GINFORS
Agricultural land use (in hectares) for 13 crop types and pastures in 38 countries and Rest of the World for the successive year	GINFORS ⇨ LPJmL
Water abstraction (in cubic metres) by the public water supply, by industry and for production of electricity [for cooling purposes] in 38 countries and Rest of the World for the actual year	GINFORS ⇨ LPJmL
Land scarcity factor for 38 countries and Rest of the World for the successive year	GINFORS ⇨ LPJmL
Total irrigated area (in hectares) for 38 countries and Rest of the World for the actual year	LPJmL ⇨ GINFORS

Table 5: Content of the data exchange between LPJ and GINFORS

The model LPJmL explains the production quantities (in tonnes) for 13 crop types (see Table 7). The calculation is executed for grid cells and then aggregated into the geographical resolution of GINFORS [38 countries and Rest of the World]. This physical information on the realised harvest in every year is transferred from LPJmL as input to GINFORS.

No.	Crop type
1	Temperate cereals (barley, rye, wheat)
2	Rice
3	Maize
4	Tropical cereals (millet, sorghum)
5	Pulses
6	Temperate roots (sugar beet)
7	Tropical roots (cassava)
8	Sunflower
9	Soya beans
10	Groundnuts
11	Rapeseed
12	Sugarcane
13	Others (e.g. fruits, vegetables, cotton, cocoa, coffee, potatoes, palm oil)

Table 7: Classification of crop types in the data exchange between LPJ and GINFORS

Within GINFORS, the evolution of crop demand for domestic purposes as well as for exports is explained. Amendments of prices for crops result in a global balance of demand and supply, for each of the 13 crop types, in physical units. On the national level, despite price amendments, domestic demand and production do not match. The balance of demand and supply is achieved by amendments to the trade balance. At the end, the evolution of prices for crops [and/or of total domestic and export demand for crops] serves to explain to farmers the decisions about agricultural land use in the next year.[3] If the price [or demand] of one crop type rises more strongly than the one for another crop type, the farmers will try to expand the area under cultivation for the type with the higher price [or demand] growth whilst diminishing the area under cultivation for the other one. The resulting agricultural land use for the successive year is transferred from GINFORS as input to LPJmL.

Another issue that is covered by GINFORS is the explanation of water abstraction by the public water supply, by industry and for the electricity supply for cooling purposes (see Deliverable D3.2). The results for the actual year are also transferred as input to LPJmL. There the evolution of water demand in combination with information about water availability leads to modifications to the irrigated area (as a share of total agricultural land use).

The next content of data exchange between the two models refers to land scarcity. As already mentioned, within GINFORS the agricultural land use in the next year is modeled bottom-up for 13 crop types and the pastures area. On the other hand, the total available land for agricultural

---

[3] This decision is also modelled with regard to the pasture area based on information about the evolution of domestic and export demand for meat and other livestock products (see Deliverable D3.2).

uses in a country/region might be restricted due to physical constraints as well as to ecological targets. As long as the aggregated area for all these purposes does not exceed this capacity, everything is fine. But in cases where the aggregated area exceeds the available capacity, the data transferred to LPJmL already contains amendments to these boundaries. To ensure that results from GINFORS with regard to the extent of land scarcities can be used in LPJmL (for amendments of the irrigated area), a land scarcity factor for each of the 38 countries and Rest of the World is calculated and is a further topic of the data exchange between the models.

One possible adaptation to an increasing (agricultural) land scarcity is an enhancement of the irrigated area. A precondition therefore is either the availability of water or the condition that the total water abstraction remains below the ecologically acceptable limit. But if this is the case and within the LPJmL model this adaptation process is implemented, the results for the total irrigated area in LPJmL have to be transferred to GINFORS to ensure that the investment needs that are induced by an expansion of the irrigated area can be implemented within the economic cycle.

3.2 LPJmL

3.2.1. Technical description of LPJmL

The Lund-Potsdam-Jena dynamic global vegetation model with managed lands (LPJmL) uses process-based representations of major bio-geochemical, bio-geographical and bio-geophysical processes to simulate the role of vegetation and soils in the Earth system, in particular with respect to their influence on the global cycles of carbon and water, the effects of human land use on the global environment, and the impacts of climate change on natural ecosystems and agriculture. Patches of plant functional types (PFTs) that coexist within a model grid cell represent natural vegetation composition at the biome level. Nine PFTs are defined by a set of functional and morphological traits describing structure (trees versus herbaceous plants), photosynthetic pathway (C3 or C4), leaf type (broad-leaved or coniferous) and phenology (evergreen versus deciduous) and other characteristics. The potential range of PFTs is constrained by a set of bio-climatic limits based on the dependence on climatic conditions of physiological processes related to plant growth and reproduction. PFTs compete for light, water and space within a single grid cell. The representation of croplands and pastures uses 13 crop functional types (CFTs), corresponding to the world's most important economic crops, and two managed grass types, respectively. The most recent model version also includes three highly productive biomass functional types (BFTs) for bioenergy production. Two tree species for temperate and tropical regions, and one fast growing grass. Tree BFTs were parameterised as temperate deciduous, to match the field performance of poplars and willows, and tropical evergreens, respectively, to reproduce the growth and biomass production of appropriate eucalyptus species. Energy trees are managed as short rotation crops and coppiced every 8 years. The energy grasses reflect the growth and productivity characteristics of miscanthus and switchgrass cultivars. In the simulations grasses are harvested annually at the end of the growing season. The calculation of photosynthesis, following the process-based bio-geochemical approach by Farquhar et al., 1980, and Collatz et al., 1992. NPP (net primary production), is the balance between carbon uptake (photosynthesis) and release (autotrophic respiration), and describes actual plant growth. Fixed carbon is then allocated to different biomass compartments (leaves, sapwood, hardwood, roots) following a set of fixed coefficients. Compartment-specific rates of tissue turnover and plant mortality determine the flux of dead biomass into the litter pool from where it is ultimately transferred into the soil carbon pool. Several soil pools are defined for each grid cell with different decomposition rates to account for labile and more resistant soil carbon pools. Soil respiration is estimated as a function of temperature and soil moisture based on a modified Arrhenius formulation in combination with an empirical soil moisture relationship. Fire is the most important disturbance factor on the global scale and is therefore also simulated by LPJmL. A macro-scale representation of the global water cycle simulates interception, soil evaporation, and the influence of snow and permafrost on the seasonality of runoff (Gerten et al., 2004) and lateral water transport between grid cells (Rost et al., 2008). A modified permafrost module was implemented recently (Schaphoff et al., 2013). Carbon and water cycles are fully coupled. LPJmL is driven by monthly fields of temperature, precipitation, cloud cover, atmospheric CO2 concentration, and land use (Sitch et al., 2003, Bondeau et al., 2007). The model has been successfully evaluated against various observational data, such as vegetation activity measured by leaf area index (Lucht et al., 2002), biosphere-atmosphere carbon exchange (Peylin et al., 2005), runoff (Gerten et al., 2004), and agricultural yields (Bondeau et al., 2007).

3.2.2. Development of LPJmL related to the model coupling

The coupling with GINFORS requires two key modifications and extensions of LPJmL. Firstly, a novel land use allocation module has to be developed that bridges the different spatial resolutions of GINFORS and LPJmL, and is thus a prerequisite to the coupling. Secondly, an extension to the input/output module will allow reading and writing of simulation data in the above-described CSV format (see Chapter 2.2) used by the models to exchange data.

New land use allocation module

LPJmL runs on 0.5°x0.5° spatial resolution while GINFORS simulates 38 countries and the combined residual category “Rest of the World”. A key challenge in the coupling is therefore the transformation of the gridded data from LPJmL into regional data for GINFORS and vice versa. The aggregation of LPJmL data into GINFORS regions is straightforward. All grid cells in LPJmL are clearly assigned to a specific country and only need to be aggregated according to the country classification in GINFORS. This applies to data on agricultural yields and irrigated areas transferred from LPJmL to GINFORS (Table 5). Agricultural yields in carbon per hectare and year will be converted into tonnes of dry matter per year using a set of CFT-specific (crop functional types) conversion factors. Using information on land use change, water demand, and land scarcity from GINFORS in LPJmL requires a spatial allocation of country-level information to the grid cells within a country. Changes in agricultural areas will be spatially disaggregated based on the suitability of the soil and climatic conditions, past land use change dynamics, economic indicators including land scarcity from GINFORS, and scenario-dependent assumptions about land availability that reflect future changes in nature conservation, biodiversity protection or land-based climate mitigation measures (e.g. afforestation). Freshwater availability for irrigation will be calculated on the watershed scale including only local renewable water resources to prevent depletion of fossil groundwater reservoirs. Irrigation is possible where excess surface runoff is available after sufficient water has been allocated to safeguard natural ecosystems. Depending on the underlying scenario, assumptions based on the POLFREE scenarios, prioritisations in freshwater use for food and biomass production, households, industry, and electricity production will be applied. Assumptions about changes in irrigation efficiency will also be adapted to the POLFREE scenarios. Simulated potentials of irrigated crop production will then be transferred to GINFORS.

New input/output module

GINFORS and LPJmL will communicate via text files with a predefined CSV format (see description in Chapter 2.2). The standard version of LPJmL uses the network-common data format (NetCDF) or the raw binary data format for input and output. An extension of LPJmL's routines that adds support of CSV to the model is therefore under development.

3.3 GINFORS

In Deliverable D3.2 a very detailed concept for new features within GINFORS has been elaborated. Although the structure of the overall modelling challenge remains unchanged, in the meantime some modifications of this concept have been developed on the basis of discussions between the modelling teams at PIK and GWS. The ambition of this Deliverable is not to repeat all of the details given in D3.2 but to introduce the main modifications to the first concept. These are:

- Within the farming module of GINFORS the classification of crop types will be aligned to the one used in LPJmL (see Table 7). Instead of distinguishing 10 crop types, the farming module will distinguish 13 crop types. By doing so, the interaction between crop production (in LPJmL) and crop demand and international trade in crops (in GINFORS) will be based on the same classification concept, and the need for an (assumption-based) adaptation of different concepts is eliminated.

- The explanation of domestic crop demand for different purposes (e.g. food, feed) will not be carried out on the basis of aggregated values for these purposes which are then distributed to the single crop types via equal shares. Instead, the explanation of crop demand in physical units will use a bottom-up approach.

- By doing so, the prices for crops and the trade balance (instead of the stock variations) serve for an alignment of supply and demand (in physical units).

- A further difference to the modelling design in D 3.2 is that GINFORS now gives the (change in) agricultural land use for the different crop types in the successive year as feed-back to LPJmL instead of "crop production aspired to". This is now a much clearer distinction. GINFORS depicts the economic decisions of the farmers on land use. The growth of crops is calculated by LPJmL.

- The modelling of the irrigated area (as a share of the total agricultural land use) for the different crop types and countries/regions is not the object of the modelling task of GINFORS but is covered within LPJmL. The mission of GINFORS is limited to delivering well-founded figures for the evolution of water demand (abstraction) by the public water supply, industry and for the electricity supply (for cooling purposes). Together with bio-physical as well as ecological/political limitations, these figures enable the model LPJmL to calculate the irrigated area.

4. Expected reactions of the models on the implemented restrictions

The inputs from the one model to the other model have the character of a restriction for the receiving system like an exogenous shock. For the success of the linking, it is useful to know the reactions that are induced. Furthermore, we have to show that the reactions to the restrictions improve the modelling results. In other words, what the advantage of the solution from the linked system is compared with a solution from the "stand alone" GINFORS.

4.1 The reactions of GINFORS to the inputs from LPJmL

GINFORS gets the production in tonnes of 13 crops for all 38 countries and the region Rest of the World. GINFORS calculates the demand for all crops in tonnes for all of the countries mentioned. We have to realise at this point that the crop demand decision is embedded in

GINFORS interdependently on the development of the global economy in deep sector and country detail including labour markets, the use of crops for feed, food and energy purposes. Furthermore, we have to mention that the direct and indirect impacts of all policy instruments, such as economic instruments (taxes, subsidies etc.) as well as regulations (for example export caps etc.), are depicted as part of the scenarios.

In every country there will be a deviation between supply and demand. Assuming that the crops are more or less homogenous goods, their prices will be identical in every country, and in this respect we can speak of a world market price for each crop. In a first step, GINFORS calculates for each crop this price that clears the world market. Implementing the world market price into the national demand functions gives the national demand, which is consistent with the world market equilibrium. The difference between the national supply and the national demand is bridged by international trade. Subtracting the price-dependent import ratio from national demand yields domestic demand. The exports of the crop in question are defined as the difference between production and domestic demand.

This is the first round effect. Let us assume that in the previous period global supply equalled global demand. If in the actual period global production is reduced by a change in plant growth ruled by natural factors, the price of the crop will rise. Concerning the crop growth, the situation in the countries will be different because the natural conditions for crop growth are different in different countries. There may be countries with rising and others – in our case dominating - with falling production, but in all countries the price for crops will rise. The latter will stabilise the income situation in those countries with falling agricultural production.

For all countries, the second round effects for the other sectors and the total economy are very similar, because the price for agricultural products rises everywhere. This effect is important since the price elasticity of demand is relatively low for agricultural products, which means that the rise in the price to clear the market is much stronger than in the case of other products. This further means that consumers have to reserve a greater part of their budget for agricultural products with the consequence of a reduced consumption demand for the other sectors of the economy. The rise of the consumer price index will induce higher wage rates and therefore higher costs in all sectors of the economy. Since demand falls in the non-agriculture sectors and the wage rate rises, unemployment will rise. These effects generate further reductions in the circular flow of income.

In developing economies, agricultural products have a much bigger share of the budget of consumers than in industrialised economies. Therefore the effects on a reduction of demand in the non-agriculture sectors of the economy and the rise in labour costs and unemployment in all sectors will be much higher in developing countries than in industrialised countries.

The general reduction in economic activity will also induce diminished water demand from the non-agricultural sectors and households, which offers a bigger amount of water for irrigation in agricultural production. This is a stabilising feedback effect into nature, because it favours the growth of plants for the next period. A second feedback effect is also stabilising the combined nature/economy system. Farmers will change their land use decisions favouring those crops with higher relative prices at the expense of the other crops. This means that the scarcity of the crops with the risen prices will be reduced.

Without the link to LPJmL, nature would have no impact on the economic system. Crop demand would be facing exogenous world market prices without any possibility to link their development with the growth conditions of plants. In particular the feedback effects would be beyond any reasoning.

4.2 Expected model behaviour

From the perspective of LPJmL, the coupling with GINFORS mainly involves a change in the source of input data driving LPJmL and not a substantial change in model structure or functioning. Therefore we do not expect numerical instabilities to result from the coupling. Information on land use and water demand from GINFORS, however, will drive land use dynamics in LPJmL and will thus have a large effect on the modelling results.

5. Conclusions and outlook

The coupled system will allow integrated quantitative analyses of the environmental consequences of policies and socio-economic development through their effects on land use and water demand. Such modelling studies are not possible with the stand-alone version of LPJmL. Trade-offs between different policy goals, such as developing a bio-based economy, increasing food security, decarbonising the energy system with bioenergy, or climate change mitigation through afforestation are also only possible in the coupled system. The role of trade in global land use displacement or changes in CO2 emissions from change in land use, driven by economic activities and resource use, are also key dynamics of the land system that we are able to study with GINFORS-LPJmL.

The logical, the technical and the data problems have been solved by now. For the future, the challenge will still be to solve the linked system. This is the case since – as already mentioned – a lot of income and price effects will be induced in GINFORS which may destabilise the system. This will not preclude a successful completion of the exercise, but it will take some time and it may be necessary to marginally change the approach.

6. References

Barton, D. N., Saloranta, T., Moe, S. J., Eggestad, H. O., & Kuikka, S. (2008). Bayesian belief networks as a meta-modelling tool in integrated river basin management: Pros and cons in evaluating nutrient abatement decisions under uncertainty in a Norwegian river basin. *Ecological Economics*, 66(1), 91–104.

Bondeau, A., Smith, P. C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., Lotze-Campen, H., Müller, C., Reichstein, M., & Smith, B. (2007). Modelling the role of agriculture for the 20th century global terrestrial carbon balance. *Global Change Biology*, 13(3), 679–706.

Braat, L. C. & van Lierop, W. F. J. (1987). Integrated economic-ecological modeling. In L. C. Braat & W. F. J. van Lierop (Eds.), *Economic-ecological modeling* (pp. 49–68). Amsterdam: North-Holland.

Brouwer, R. & Hofkes, M. (2008). Integrated hydro-economic modelling: Approaches, key issues and future research directions. *Ecological Economics*, 66, 16–22.

Cai, X., Ringler, C., & You, J.-Y. (2008). Substitution between water and other agricultural inputs: Implications for water conservation in a river basin context. *Ecological Economics*, 66(1), 38–50.

Collatz, G., Ribas-Carbo, M., & Berry, J. (1992). Coupled photosynthesis-stomatal conductance model for leaves of C_4 plants. *Australian Journal of Plant Physiology*, 19(5), 519–538.

Farquhar, G. D., Caemmerer, S. V., & Berry, J. A. (1980). A biochemical model of photosynthetic CO_2 assimilation in leaves of C3 species. *Planta*, 149(2), 78–90.

GAMS Interfaces Wiki (2014). Gams return codes. http://interfaces.gams.com/doku.php?id=env:gams_return_codes.

Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W., & Sitch, S. (2004). Terrestrial vegetation and water balance-hydrological evaluation of a dynamic global vegetation model. *Journal of Hydrology*, 286(1-4), 249–270.

Jonkman, S. N., Bockarjova, M., Kok, M., & Bernardini, P. (2008). Integrated hydrodynamic and economic modelling of flood damage in the Netherlands. *Ecological Economics*, 66(1), 77–90.

Lucht, W., Prentice, I. C., Myneni, R. B., Sitch, S., Friedlingstein, P., Cramer, W., Bousquet, P., Buermann, W., & Smith, B. (2002). Climatic control of the high-latitude vegetation greening trend and Pinatubo effect. *Science*, 296(5573), 1687–1689.

McKinney, D., Cai, X., Rosegrant, M. W., Ringler, C., & Scott, C. A. (1999). Modeling water resources management at the basin level: Review and future directions. *International Water management Institute, Colombo: SWIM Paper*, 6.

Peylin, P., Bousquet, P., Le Quéré, C., Sitch, S., Friedlingstein, P., McKinley, G., Gruber, N., Rayner, P., & Ciais, P. (2005). Multiple constraints on regional CO_2 flux variations over land and oceans. *Global Biogeochemical Cycles*, 19, GB1011.

Pulido-Velazquez, M., Andreu, J., Sahuquillo, A., & Pulido-Velazquez, D. (2008). Hydro-economic river basin modelling: The application of a holistic surface-groundwater model to assess opportunity costs of water use in Spain. *Ecological Economics*, 66(1), 51–65.

Rosenthal, R. E. (2013). Gams - A user's guide. http://www.gams.com/dd/docs/bigdocs/GAMSUsersGuide.pdf, GAMS Development Corporation, Washington D.C.

Rost, S., Gerten, D., Bondeau, A., Lucht, W., Rohwer, J., & Schaphoff, S. (2008). Agricultural green and blue water consumption and its influence on the global water system. *Water Resources Research*, 44, W09405.

Schaphoff, S., Heyder, U., Ostberg, S., Gerten, D., Heinke, J., & Lucht, W. (2013). Contribution of permafrost soils to the global carbon budget. *Environmental Research Letters*, 8(1), 014026.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., & Venevsky, S. (2003). Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, 9(2), 161–185.

Volk, M., Hirschfeld, J., Dehnhardt, A., Schmidt, G., Bohn, C., Liersch, S., & Gassman, P. W. (2008). Integrated ecological-economic modelling of water pollution abatement management options in the Upper Ems River Basin. *Ecological Economics*, 66(1), 66–76.

Ward, F. A. & Pulido-Velázquez, M. (2008). Efficiency, equity, and sustainability in a water quantity - quality optimization model in the Rio Grande basin. *Ecological Economics*, 66(1), 23–37.