Plagiarism without Apology – Systematic Integration of Available Information in a Long Run Agricultural Outlook

H. Peter Witzke
Wolfgang Britz

**Paper prepared for presentation at the XIth Congress of the EAAE (European Association of Agricultural Economists),
'The Future of Rural Europe in the Global Agri-Food System',
Copenhagen, Denmark, August 24-27, 2005**

PLAGIARISM WITHOUT APOLOGY – SYSTEMATIC INTEGRATION OF AVAILABLE INFORMATION IN A LONG RUN AGRICULTURAL OUTLOOK

Abstract

In the context of a long run agricultural outlook on behalf of the European Environmental Agency a new methodology has been developed to systematically integrate external forecasts into a quite detailed agricultural sector model. External forecasts usually provide estimates for the exogenous variables in modelling work and frequently they are also used for comparisons and potential reassessment of empirical specifications. The innovative characteristic of this study is that expert forecasts have been used to specify parameter changes expressing structural change affecting behavioural functions. The outlook was therefore set up as a simultaneous estimation and forecasting effort which permitted to integrate various, usually contradictory expert forecasts subject to the equations of the sector model.

Keywords:

Agricultural outlook, forecasting, modelling, expert information

JEL classification: C15, C53, Q11, Q19, Q21

1. Introduction

This paper will present the procedure applied in the context of an agricultural outlook on behalf of the European Environmental Agency. There were quite demanding requirements for this outlook: A high level of detail, sufficient to calculate environmental indicators, was to be combined with a very long run time horizon (up to 2025). Furthermore it was required that the analysis should not merely add another projection to the available set of outlooks by other agencies but that the study should systematically build upon the existing knowledge.

To capture key driving forces the projection relied on a quite detailed agricultural sector model (CAPSIM) with behavioural functions for key endogenous variables based on microeconomic theory. However, as certain long run developments such as farm structure changes could not be depicted in explicit form it was clear that expert information had to supplement the formal framework. External forecasts usually provide estimates for the *exogenous* variables in modelling work. The innovative characteristic of this study is that expert forecasts have also been used to specify selected model parameters related to *endogenous* variables which might be affected by structural change of any kind. The outlook was therefore set up as a simultaneous estimation and forecasting effort which permitted to integrate various, usually contradictory expert forecasts subject to the equations of the sector model. Due to various interrelationships the approach required to solve the model simultaneously for 22 EU Member States (Belgium and Luxembourg were combined, Malta and Cyprus omitted due to data limitations) and three explicit simulation years (1994, 2011, 2025) with an implicit base year (2001). Given the high level of detail in products and activities, this resulted in about 45.000 equations and 52.000 variables and thus imposed technical constraints on feasible procedures.

The paper will be orgnised as follows. The next section will discuss the background of the approach in more detail. The third section will give a brief overview on the default trends supplementing the expert forecasts as an external source of information. The fourth section highlights the key relationships of CAPSIM which are prerequisites for the 'integration task' of the study. The fifth section directly focusses on the integration methodology and presents selected examples. The sixth section briefly reports on the main results of the agricultural outlook and the final section concludes.

2. Background and problem definition

The application of a quantitative model for impact analysis requires first to specify a reference run or 'baseline' which serves as a yardstick for subsequent identification of impacts . The baseline is usually presented cautiously in the form of conditional forecasts, as a consistent set of predictions which critically depend on a set of key assumptions, chosen to represent plausible developments. Performing a reasonable reference run is a forecasting or 'outlook' task tackled in various organisations such as the FAO (Bruinsma 2003), OECD (Conforti, Londero 2001), FAPRI (Westhoff, Young 2001: 257) or DG Agri (e.g. EU Commission 2004). Usually it relies on some form of econometrics from simple trends to advanced time series techniques. Typically the forecasts are not produced by a single overall coherent model but based on a set of interlinked models tailored to analyse specific agricultural subsectors ('cereals', 'oilseeds', ...). A final characteristic of serious outlook work is that information from many market observers and participants is integrated in an informal way, either from market or regional experts within large public agencies (FAO, DG Agri) or from outside in organised discussions meetings (FAPRI, AGLINK process). This informal integration of information may pick up ongoing 'structural' changes within agriculture (farm size distribution, part time farming, technological progress...) and the food sector (consumer tastes, marketing margins) which are already difficult to capture in small case studies but impossible to tackle for a differentiated set of products and many regions in a limited period of time.

Whereas a good fit to past data is an important quality criterion for pure outlook work policy simulations require to work out endogenously the consequences of modified, potentially new policy instruments, in a transparent form . Transparency is more important in policy simulations than in outlook work because a resonable model structure may be the only secure basis to build on. Especially if the simulated policy represents a significant change to the status quo, the experience of market observers is of limited value if nothing comparable has been observed in the past. Econometric estimations may be a shaky basis as well, if the structure changes as a consequence of the policy change (Lucas 1976). This may apply as well to trend shifts if they are capturing structural changes which may respond to large policy changes. However it may also apply to other parameters deemed invariant over scenarios such as price elasticities. Each modelling effort has to be accompanied therefore by a critical judgement on which presumably exogenous variables or parameters might change in spite of assumed exogeneity.

Of course, the distinction of modelling systems for outlook work and for policy simulations is not a strict one. CAPSIM has a focus on policy simulations but in view of the necessary reference run it has been prepared to integrate various pieces of information from different sources in a systematic way. To serve these two purposes the system is applied in two "modes": The *reference run mode* is used to calibrate certain time dependent parameters (shifters) in model equations, building on exogenous forecasts and ex post observations for the related variables, for example the activity levels. The *policy simulation mode* of CAPSIM utilises these calibrated values for parameters to simulate the impacts of alternative policies and exogenous inputs with given parameters. For those exogenous inputs, which are not open to modification, the standard rule is that they will be taken over from the reference run, for example yields, final consumption expenditure, and the inflation rate. When policy inputs and other inputs are chosen equal to their reference run values in the policy simulation mode the outcome reproduces exactly the reference run results.

The reference run mode resembles the AGLINK approach where so called "add factors" shift the model results to match forcasts from OECD Members given assumed prices (Conforti, Londero 2001). The same occurs in the calibration of FAO-@2030 to a new baseline (Britz 2003: 46). As DG Agri obtained quite consistent results in their recent simulations for the "Prospects" publication using ESIM (Münch 2002) and the in house "set of partial equilibrium, dynamic models" (EU Commission 2003:

10) it may be speculated that some shifters have been used to ensure this consistency. Apart from these examples of systematic calibration of reference run results to an existing set of forecasts it is quite clear that each large scale modelling system includes thousands of parameters and that some of them will be reconsidered in case of surprising results. The revision may be made in a reestimation or on the basis of "expert judgement" (Westhoff, Young 2001: 257) which is just another example of parameter adjustments being used to obtain a plausible baseline.

In the outlook work on behalf of the EEA, these parameter adjustments for the reference run have been placed on a systematic footing:

- CAPSIM will internally merge a set of external forecasts even if they contradict each other and if they are technically infeasible.
- In addition to the model's base year another ex post observation is included to add information for the development of time dependent parameters.

In essence the expert forecasts are treated as if they were ex post observations and the development of shifters is chosen to maximise the "fit" of model outcomes compared to observed or forecasted information. In several cases, in particular for 'unimportant' (in the EU say textile crops) or 'difficult' items (say fodder production) there were no expert projections at all. In particular for these cases it was necessary to prepare a set of 'default' trends.

3. Default trends with built in safeguards

Given that it was impossible to know in advance for which agricultural variables default trends would be needed estimates were prepared for all of about 1500 time series per Member State. In view of their later use for very long run projections several safeguards against unreasonable projections have been built into them.

The first safeguard is a trend function ($a+b\ t^c$, $0 \le c \le 1.2$) restricting the maximum change over time. The second is a two step procedure where the forecast of a badly fitted series would be pulled towards the most recent base year data. Note that in a standard forecasting effort the trend term would be ignored if the t-value of the associated trend parameter did not exceed conventional significance levels. We applied this rule in a continous form rather than with a threshold significance level (noting the relationship of R^2 to the t value in simple regressions) when defining a 'target' value for the variable in question as

$$X_{r,j,t}^{i,target} = R_{r,j,i}^{2} X_{r,j,t}^{i,step1} + \left(1 - R_{r,j,i}^{2}\right) X_{r,j,bas}^{i,data} \qquad (1)$$

where the 'step1' forecast is the standard ex ante trend forecast based on 1985-2002 data for product j, item i, region r in a forcast year t, the base year 'data' represent the five year average 1998-2002, and the resulting weighted average is the 'target' value for step 2 of our trend estimation methodology. Weighting the step 1 forecasts with R^2 tends to produce conservative estimates close to the base year value if the fit was bad.

Step 2 introduces a third group of safeguards. This is a quite exhaustive set of technological constraints and identities tying together the series. If we ran independent trends for, say, barley area, yield and production the forecasts would almost certainly violate the identity linking the three series. One solution would be to drop one of the three from the estimation, for example the yield, and to compute it later from the forecasts of the two other series. However, by doing so, we would ignore the information incorporated in yield observations which is avoided in our simultaneous approach. As a consequence, we have imposed several balances (for feed energy and crude protein, milk fat and protein in dairies, markets including land and nontradable young animals) and identities (production = activity levels * yields, values = quantity * price, consumption = per capita consumption * population, aggregates = sum of components). Furthermore for several variables assumed to change only slowly, maximum yearly growth rates have been defined, e.g. for total agricultural area (max ± 0.5% per year), cereal consumption per head (max ± 0.4%), and so forth. The (final) step 2 forecasts are those values which minimise normalised squared deviations from the 'targets' while meeting all the above constraints:

$$\mathrm{obj} = \sum_{r,i,j,t} \left(\frac{X_{r,j,t}^{j,step2} - X_{r,i,t}^{j,target}}{S_{r,j,i}^{step1}} \right)^2 \qquad (2)$$

where we used the standard errors of each independent step 1 trend to normalise the deviations. This normalisation yields contributions to the objective which are independent of the units of variables but will be scaled down in case of a bad fit in step 1.

Finally we may note here that for a single but very important variable for the lifestock sector it was easy to incorporate some knowledge about agricultural policy: In line with the reference run assumptions of this study cow milk production has been fixed at the quota levels assumed unchanged after 2015.

4. Key relationships in CAPSIM

The default trends and expert predictions are merged in the framework of the agricultural sector model CAPSIM. It is a straightforward partial equilibrium modelling tool with behavioural functions for activity levels, input demand, consumer demand and processing. The *product list* includes 30 marketable agricultural products, 5 nonmarketable agricultural products (fodder maize, other fodder, grass, male calves, female calves), 17 processed products, and 3 rather aggregate inputs (general cost items, plant related inputs, primary factor aggregate). It is also quite detailed regarding CAP policy instruments and covers individual EU Member States. Complete coverage of the whole of agriculture permits to include certain technical relationships, e.g. adding up of total areas or balancing of feed contents and animal requirements. Firm microtheoretic underpinnings are considered useful to increase the plausibility of simulation behaviour, and to understand the model's properties which is particularly useful for long run analysis.

4.1 Supply side

As in other acricultural sector models (mentioned above) production is given by the product of activity levels and yields. In the context of this study this permitted to calculate simplified environmental indicators. Yields have been specified as unresponsive to prices which is admittedly a simplification. However it appears that the major part of supply responsiveness in the crop sector is coming from ara reallocations rather than from yield responses such that the exogenous yields simplification may be defendible (also used in CWFS/CPB 2003: 149, for example).

If yields are exogenous from the farmers point of view, then so will be revenues per activity which may be combined therefore with input prices to a set of exogenous "price variables" $X_{r,i,t}^{PVAR}$ (in region r, for product i, in year t) as determinants of so called "netputs" combining both activity levels ($X_{r,i,t}^{NETP} > 0$) and input demands ($X_{r,i,t}^{NETP} < 0$).

$$X_{r,i,t}^{NETP}(\boldsymbol{X}_{r,t}^{PVAR}) = \alpha_{r,i,0,t}^{NETP} + \sum_j \alpha_{r,i,j}^{NETP} X_{r,j,t}^{PVAR} \qquad (3)$$

These behavioural functions derive from Normalised Quadratic profit functions calibrated to meet microeconomic consistency and to reflect plausible assumptions on elasticities. Balances of land, feed energy and crude protein are imposed through a shadow price adjustment of the price variables (Witzke 2005). Note that the constants carry a time index and may hence be shifted over time, in this functional form without affecting symmetry or convexity of the underlying profit function. These shifts may be due to technological progress (beyond yield growth), changes in farm structure, fixed factor endowments (labour, capital) or any other changes over time which are not captured by the price variables. In view of the long run projection horizon it would have been desirable, of course, to model these determinants explicitly and endogenously but given the requirement to cover agriculture in a high level of detail this was infeasible in the available time. These limitations may be compensated, however, through expert based shifts of behavioural functions.

In a standard econometric study some parametric form would be postulated, say a linear or logarithmic trend for $\alpha^{NETP}_{r,i,0,t}$, to proceed to an estimation of the yearly shift. The constant terms in this study are also specified to maximise a 'best fit' criterion, but the 'observations' to be represented are in part future projections from external sources. For this 'estimation' or 'multi year calibration' problem it turned out useful to have more flexibility than offered by common transformations while avoiding computational complexity as in the Box Cox form. Furthermore with about 600 activity level equations it was considered prudent to impose that the shifts be monotonic and at most linear (falling or increasing) because manual checking would have been too time consuming. These goals could be achieved with spline trend functions restricted by equation (4),

$$\alpha^{NETP}_{r,i,0,t} - \alpha^{NETP}_{r,i,0,t-1} = mon^{NETP}_{r,i,t}\left(\alpha^{NETP}_{r,i,0,t-1} - \alpha^{NETP}_{r,i,0,t-2}\right) \qquad (4)$$

where the monotonicity parameter *mon* is restricted to the range [mon_{lo}, 1], and typically mon_{lo} = 0.1. An example for the implied restrictions on the spline trends is given in **Figure 1**.

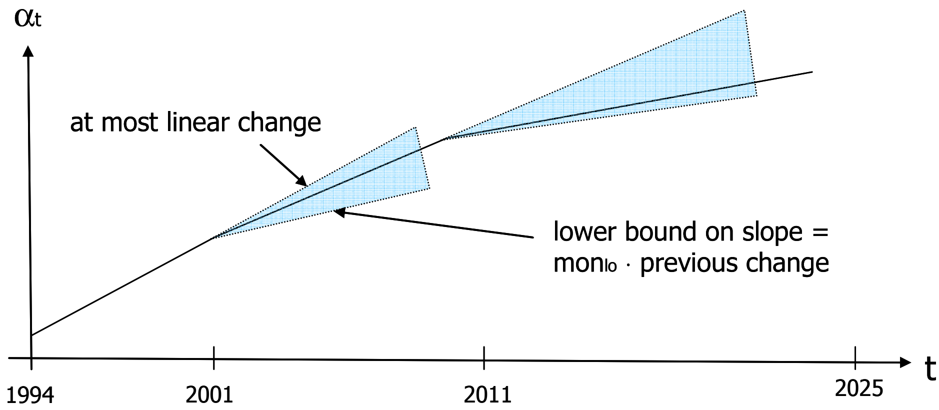

Figure 1: Example spline trend line for a constant term

The monotonicity restriction has been imposed on the constants of behavioural functions for netputs, processing demand and human consumption. It renders it difficult for CAPSIM to track non-monotonic expert forecasts of, say, an activity level over time, if this development is not due to price changes. Equally CAPSIM will not be able to trace an geometrically growing series. The idea incorporated in this way is that any growth found in recent trends is more likely to taper off than to continue forever or even to explode, at least if we have to ignore short run stochastic influences and focus on the mean expectation. Not all variables will really shift monotonically with tempered trends in the future. However, it is expected that the restriction increases the plausibility of projections for the majority of variables.

The constraining character of the trend lines for shifters may be recognised easily if we consider an example. Assume that the 1994 ex post value for a variable was 110, the 2001 base year value was 100 and for 2011 we had a single credible external projection of 150 (Ignore the fourth simulation year 2025 for the moment). All price changes are assumed to cancel in their effect such that the sum of all price related terms in the behavioural function (3) is 80 in each year. Free constants would take on values of $\alpha_{r,i,0,1994}$ = 30, $\alpha_{r,i,0,2001}$ = 20, $\alpha_{r,i,0,2011}$ = 70 to perfectly reproduce the 'observations'.

However, the sequence 30, 20, 70 would represent a non-monotonic development violating constraint (4) above. The model would look for a monotonic compromise sequence, say 15, 20, 60, which implies some deviations from the 'observations'.

Because the price related parameters of (3) are calibrated based on value shares of the base year (= three year average 2000/2002) it was considered useful to define the constant term $\alpha^{NETP}_{r,i,0,t}$ at the base year point such that the model reproduces the base year situation exactly:

$$X^{NETP}_{r,i,t_{bas}}(\mathbf{X}^{PVAR}_{r,t_{bas}}) = \alpha^{NETP}_{r,i,0,t_{bas}} + \sum_j \alpha^{NETP}_{r,i,j} X^{PVAR}_{r,j,t_{bas}} \tag{5}$$

To summarise: The key driving forces operating on the supply side are technology shifts in a wide interpretation, yield developments and price variables. Depending on the trade regime the latter are determined from the interaction with the demand side or derived from international projections.

Policy is strongly modifying the incentives on the supply side. Gross revenues of activities stem from market revenues and different types of premiums. The decoupled MTR premiums are incorporated as uniform premiums for eligible crops which appeared to be a reasonable long run assumption. Obligatory set aside and non food production are specified according to EU Commission 2004. The milk quota nd sugar regimes are assumed to remain in place until 2025, but for sugar a price and quota cut has been implemented.

4.2 Demand side

The main endogenous domestic demand components apart from feed demand are processing of agricultural raw products and human consumption, where feed demand is included in the supply side behavioural functions already. Processing demand $X^{PRCM}_{r,i,t}$ results from linear behavioural functions of unit margins $X^{UMAP}_{r,i,t}$ in processing, calculated as the difference of the value of derived products and the associated raw product cost:

$$X^{PRCM}_{r,i,t}(\boldsymbol{X}^{UMAP}_{r,t}) = \alpha^{UMAP}_{r,i,0,t} + \sum_j \alpha^{PRCM}_{r,i,j} X^{UMAP}_{r,j,t} \tag{6}$$

As is the case for the supply side behavioural equations we permitted the constant terms to shift over time along smooth monotonic trend functions. Production of secondary products is derived from processed raw products with (base period) processing coefficients which also enter the calculation of raw product cost.

A special case relates to milk products. Similar to the case for the feed technology it is considered useful to explicitly control the balances on milk fat and protein. This is achieved through associated endogenous prices of milk fat and protein. Production of secondary milk products responds to margins calculated as secondary product value net of milk fat and protein cost. Otherwise the behavioural functions correspond to the standard case in

$$X^{PRCM}_{r,i,t}(\boldsymbol{X}^{UMAP}_{r,t}) = \alpha^{UMAP}_{r,i,0,t} + \sum_j \alpha^{PRCM}_{r,i,j} X^{UMAP}_{r,j,t} \tag{6}$$

Human consumption $X^{HCOM}_{r,i,t}$ is determined according to the rather robust "Linear expenditure system":

$$X^{HCOM}_{r,i,t}(X^{EXPD}_{r,t}, \boldsymbol{X}^{UVAD}_{r,t}) = X^{INHA}_{r,t} \cdot \left(\alpha^{HCOM}_{r,i,0,t} + \frac{\alpha^{HCOM}_{r,i,1,t}}{X^{UVAD}_{r,i,t}} \left(X^{EXPD}_{r,t} - \sum_j \alpha^{HCOM}_{r,j,0,t} X^{UVAD}_{r,j,t} \right) \right) \tag{7}$$

as a function of total final consumer expenditure $X_{r,t}^{EXPD}$, inhabitants $X_{r,t}^{INHA}$, the vector of consumer prices (calculated as unit valus) $\mathbf{X}_{r,t}^{UVAD}$, and "commitments" $\alpha_{r,i,0,t}^{HCOM}$ which may be interpreted for food items quite neatly as price independent subsistence consumption levels. In case of human consumptions these commitments took over the role of time varying constant terms, that is they where allowed to shift along monotonic spline trend functions. In contrast to the endogenous commitment changes we incorporated the well known tenbdency of marginal budget shares to decrease over time (compare Lampe 1999, Rosegrant et al. 2001) in an exogenous manner.

Demand and supply side interact on markets. For tradable products international prices (border prices) are linked to EU prices using a price transmission equation based on the law of one price. Without border measures, these international prices would directly apply to EU markets. Price policy instruments are *tariffs* or, until tariffication is complete, *administered prices* with associated *flexible levies* or *export subsidies*. For nontradable products (fodder, calves) market clearing occurs on the level of Member States.

5. Integration of external projections

5.1 Formal framework

Section 4 of this paper has presented the options created to shift behavioural functions over time. In this section we will explain how this flexibility is used. For the modelling efforts related to the 2005 'State of the Environment and Outlook Report' it was decided to systematically build upon the existing projections and thereby to find an optimal compromise between possibly conflicting pieces of available information (different expert forecasts, trend forecasts and the model structure), rather than merely adding another projection. A well known interpretation of an 'optimal' compromise between distinct pieces of a priori information is provided by a cross entropy approach (Golan, Judge, Miller 1996). However, the cross entropy approach turned out impractical in this large scale modelling effort, because it introduces for each variable of interest auxiliary variables (probabilities) and equations (probability adding up, posteriori mean of supports). In this study we used an attractive and computationally less demanding alternative, the 'Highest Posterior Density' (HPD) estimator (Heckelei, Mittelhammer, Britz 2005). Nonetheless it was useful to organise the external forecasts fom source s (expert forecasts, trend forecasts) in the format of *supports* $Sup_{r,i,s,t}^{j}$ for region r, product i, source s, item j, year t with associated *a-priori probabilities* $\mathrm{pr}_{r,i,s,t}^{j}$. Usually all expert sources carry the same probability weight, except for the trend support, which received only half of the standard probability. This should reflect the fact that in spite of considerable technical sophistication (see section 3) the trends are ignorant about many non trending developments of crucial determinants (for example the latest CAP reform package). In rare cases, for example regarding projections on the rye market, the probability weight for a particular expert source (here: DG Agri) has been increased, assuming that it is particularly well informed on this isssue. The supports and a priori probabilities were used to compute the mean and standard deviation oft the a-priori probability distribution which was assumed to be normal:

$$pd = \prod_{r,i,j,t} \frac{1}{\sqrt{2\pi {\sigma_{r,i,t}^{j}}^{2}}} e^{\left(\frac{X_{r,i,t}^{j} - \overline{X}_{r,i,t}^{j}}{\sigma_{r,i,t}^{j}}\right)^{2}} \tag{8}$$

where we have assumed a diagonal covariance matrix with standard deviations,

$$\sigma^{j}_{r,i,t} = \sqrt{\sum_{s} pr^{j}_{r,i,s,t}\left(Sup^{j}_{r,i,s,t} - \overline{X}^{j}_{r,i,t}\right)^{2}} \tag{9}$$

and a-priori expected values of the variables,

$$\overline{X}^{j}_{r,i,t} = \sum_{s} pr^{j}_{r,i,s,t} Sup^{j}_{r,i,s,t} \tag{10}$$

to summarise the information included in the originally discrete distribution of the supports. The implicit normal distribution does not capture skewness in the original distribution of the supports but the main reason for skewness, namely a lower bound of zero for many variables, may also be handled with a hard lower bound for the corresponding variables. In a test simulation it turned out that the cross entropy and posterior density objectives did not differ a lot in the final results which was to be expected (Heckelei, Mittelhammer, Britz 2005).

The previous equation presume that there is a positive variance which is ensured if we add two "outer" supports characterising a very high and very low value of the variable which receive only a small probability weight. These probabilities have been calculated to standardise the implied punishment for a deviation from the mean support amounting to 50% of the conceivable range.

In the implementation we took logs of (6) to obtain a convenient quadratic objective. Finally it turned out also useful to introduce additional weights in the objective characterising the importance of the variable in question, because some variables are deemed more important, say soft wheat area in France, than others, say the sheep and goat herd in Finland. "Importance" has been measured both with the (quantity) share in EU totals as well as with shares in the (monetary) national totals which have been combined with equal weights.

$$obj = \sum_{r,i,j,t} obwgt^{j}_{r,i,t} \cdot \left(\frac{X^{j}_{r,i,t} - \overline{X}^{j}_{r,i,t}}{\sigma^{j}_{r,i,t}}\right)^{2} \tag{11}$$

5.2 Implementation

In terms of expert information this study is indepted to Rosegrant et al. 2001, Bruinsma 2003, FAPRI 2003, and EU Commission 2004 which all made their projections available in electronic form. At least some of these sources offered information on the following items:

1. Activity levels (mainly for crops),
2. Yields (mainly for crops),
3. Production,
4. Demand (total, human consumption, feed, processing),
5. Trade (net trade, imports, exports),
6. Prices (EU prices, world market prices).

Other variables are largely depending on these key variables. Producer prices on the Member State level, for example, follow in most cases from the EU prices through a fixed factor of proportionality. Income may be calculated once prices and quantities are known. In this manner the closed sector model framework helps to conveniently complete the quantitative prediction in line with the predictions on key variables.

In general, change indices relative to the CAPSIM base year 2000/2002 have been calculated as follows from the expert forecasts:

$$Ind^{j}_{r,i,s,t} = \left(X^{j}_{r,i,s,t} / X^{j}_{r,i,s,t_s}\right) \cdot \left(X^{j}_{r,i,CAPSIM,base} / X^{j}_{r,i,CAPSIM,t_s}\right) \tag{12}$$

where $Ind^{j}_{r,i,s,t}$ is the change index for region r, item j, product i, year t according to source s, $X^{j}_{r,i,s,t}$ and X^{j}_{r,i,s,t_s} are levels of the corresponding variable in year t and the base year t_s of source s, respectively, and $X^{j}_{r,i,CAPSIM,base}$ and $X^{j}_{r,i,CAPSIM,t_s}$ are the levels of the corresponding variable according to the CAPSIM database in the CAPSIM base year and in the base year of source s. The second factor could be omitted for FAPRI and FAO data because we could choose the base year in the source data identical to those in our database ($t_s = base$).

In many cases the CAPSIM database is more differentiated in terms of products than the expert sources. The latter may include, for example, a projection for "coarse grains" but not for individual cereals. In this and other cases we have applied the same change factor to each member of the group. In a similar fashion we have applied the change factors for EU15 or EU 08 to each Member State of the respective group, where necessary. This procedure introduces additional assumptions beyond those from the expert sources such that our supports are only *derived from* them.

It might be considered preferable to avoid additional assumptions and to introduce supports only for those aggregates which are directly available from the expert sources. This would certainly increase transparency in the processing of supports, but it has two disadvantages. The first and more important one is that very disparate developments are almost sure to be projected for EU Member States with EU supports only. In such cases the solver will find it optimal to attain an EU target growth rate through high growth in a few Member States with little or even negative growth in others. While this may happen occasionally in the real world we would like to avoid such diversity in the results without good explanations. Note that diversity is not precluded in our current approach as the common growth assumption defines only the target values while actual simulation results may be quite heterogeneous nonetheless. The second reason is that additional equations would be necessary to aggregate the variables to the level which corresponds closely to the corresponding source.

There are a few details to be added which are specific to each of the 6 groups of items above.

1. Expert information on activity levels is available in most sources only for important crops (or crop groups), but not for animals.
2. Yield projections are also available for crops from different expert sources. They have been used as additional supports for future years whereas yields in the ex post year for 1994 were fixed.
3. Production forecasts are given for most products but they have only been used for animal products because the crop sector is sufficiently described with activity levels and yields.
4. Expert information on demand has been introduced for the major components: feed use, human consumption and processing. Where only a larger aggregate has been available (typical case: human consumption + feed use for oilseeds) we applied the change factor for the aggregate to each of the components.
5. A change index for net trade according to Equation (11) would significantly depend on the initial net trade situation which differs in each of the sources. Therefore we derived the change index for source s from the changes on the supply and demand side with weights according to the CAPSIM database.
6. Price information, in particular on world market prices in US $, currently relies exclusively on FAPRI and IFPRI projections. The exchange rate assumption has been specified as in the DG Agri projections (1.1 $/€) from 2011 onwards (modified in a sensitivity analysis). The supports for EU prices have been derived from the international prices, taking into account a possibly reduced border protection as in the case of rice.

Next we will look at the EU23 aggregate of inner supports and final projections for beef production (Figure 2). This illustrates several characteristics of the forecasting approach chosen:

- The simulation result may deviate from the a-priori mean of the external forecasts. This happens, if certain constraints prevent an exact "fit" or if this deviation helps to bring another related variable, say the suckler cow herd, closer to the mean of its supports. In this case it is likely that the calves balances built both into CAPSIM and in the default trend estimation, pulls down the beef production given a reduced supply from dairy cows and a limited responsiveness of the suckler cow herd.

- To save equations and variables the base year result is not simulated but the parameters of the model are calibrated to impose an exact fit. Introducing the expert information in the form of change indices relative to the base year also ensures a match with the historical base year observation.
- The number of supports may differ from year to year and from variable to variable according to the time horizon of the different external forecasts and their product/item coverage.
- Occasionally the supports may differ a lot. In this case the variance of the supports will be large and deviations are punished only moderately in the objective function. This has certainly contributed in our example to a forecast moving away from the mean of the expert supports.

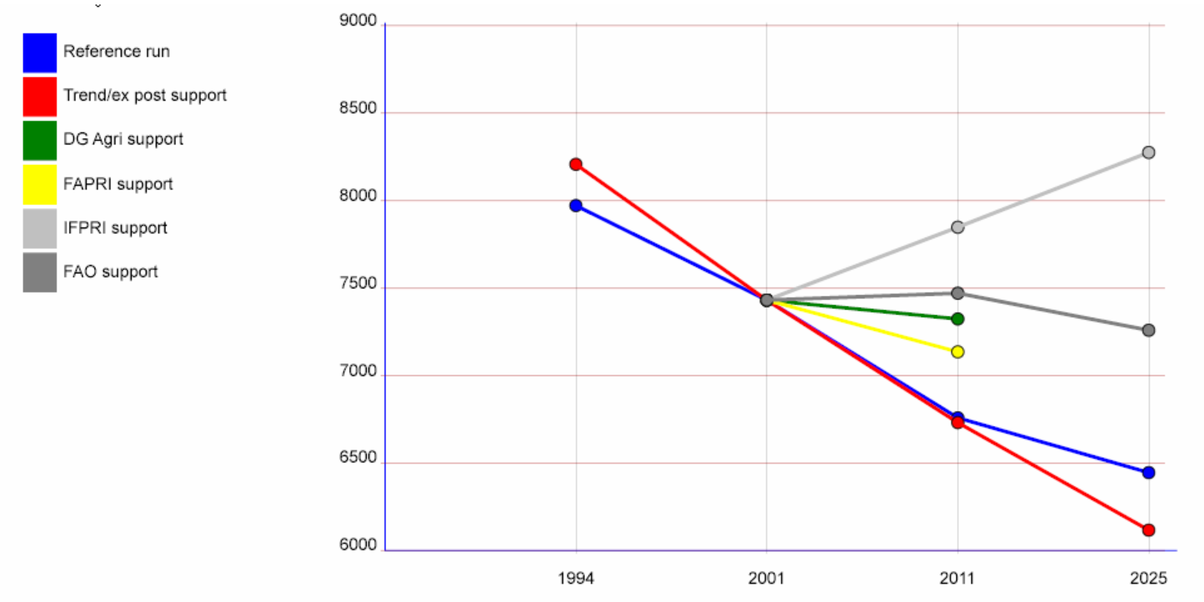

Figure 2: Aggregated "supports" and simulation results for beef production in EU23 [1000 t]

The next example has been chosen in view of the monotonicity restrictions for the behaviour of shifters. It may be seen in Figure 3 that CAPSIM may track non monotonic developments if they are explained by price movements. Feed use of potateoes in EU-23 is strongly influenced by developments in Poland where potato prices were strongly increasing prior to EU accession. As a consequence feed use in Poland strongly declined during that period which carries over to the statistical aggregate of EU-23. Between 2001 and 2011 potato prices are expected to stagnate in nominal terms such that feed demand increases again in Poland. Price movements may be another reason therefore why final projections deviate from the weighted mean of the supports, which would be in this case about halfway between the FAO and IFPRI supports, given that the default trends receive only 50% of the standard weight (see section 5.1).

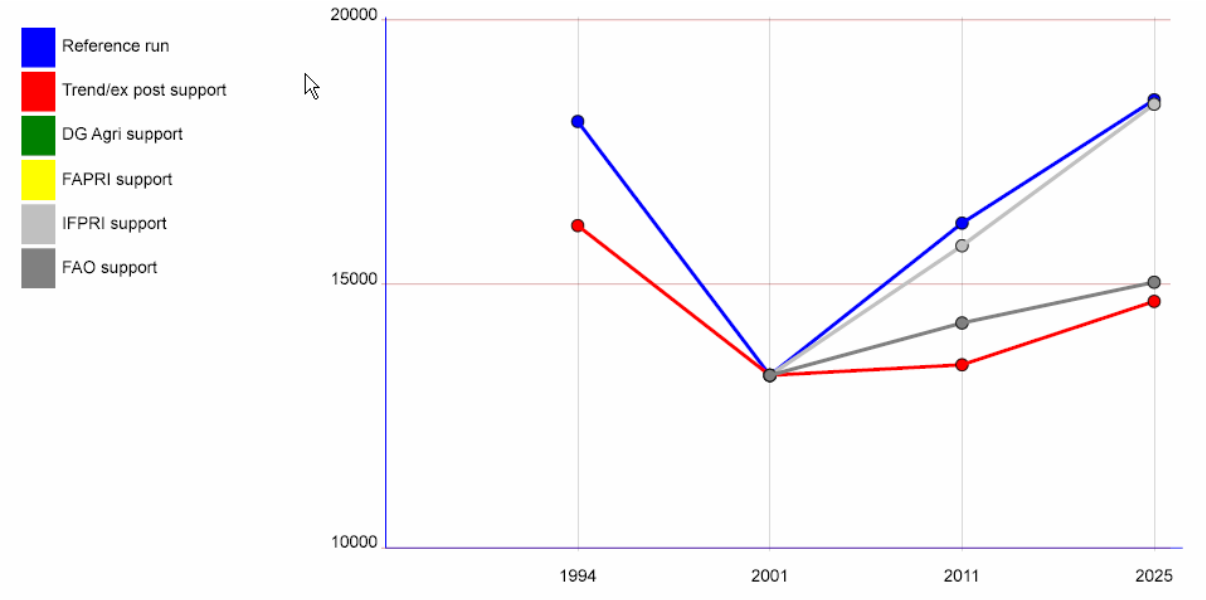

Figure 3: Simulation results and supports for feed use of potatoes in EU 23 [1000 t]

6. Overview on the substantive results of the outlook

The total area under *cereals*, during the last decade to a larger extent modulated by changes in the set-aside regime, is forecasted to fall slightly from current levels to 2011, and to recover again thereafter. These changes are more pronounced in the new Member States of the EU. The recovery is due to rather optimistic price forecasts in world cereals markets in our FAPRI and IFPRI based EU border price projections. Compared to cereals, *oilseeds* account for a much smaller part of the arable crop area. In line with DG Agri projections, oilseeds area is forecasted to continue its past expansion, almost exclusively due to rape. As markets are saturated and yields continue to grow, *potato* areas are expected to follow their long term decrease starting of from 2.7 Mio ha in 2001 to 2.3 Mio ha in 2025. Projected developments in the *sugar-beet* sector, the other major root crop, are determined by the combination of yield increases and an assumed sugar market reform involving both price and quota cuts all of which contribute to the projected drop from 2.4 Mio ha in 2001 to 1.9 Mio ha left in 2025. *Fodder area* accounts for the biggest part of the agricultural land base in EU23, with around 55 Mio ha of permanent grass land and pasture in 2001 plus another 20.8 Mio ha of arable land used for fodder maize, silage, clover etc. Fodder area shows a long term falling trend which is projected to continue, thus accounting statistically for the major share of agricultural land being lost for nonagricultural uses. About a quarter of that change would occur in the New Member States. The development in fodder areas is linked to reduced fodder demand from ruminants as both supply of beef meat and cow numbers are projected to drop in the long term. Changes in *other crops* are somewhat heterogeneous but are expected to compensate to a large extent. *Set-aside* obligations are assumed to prevail over the projection period. Structural change in the new Member States will shift Grandes cultures areas from small farms, exempt from set-aside, to larger ones which is assumed to double set-aside obligations in EU8 until 2025. Voluntary set-aside is also projected to contribute to the increase shown in Table 1 whereas (unpaid) *fallow land* would be quite stable in the EU-23 as increases in EU-08 and decreases in EU-15 cancel.

Table 1: Reference run development for crop levels

	1994	2001	2011	2015	2020	2025
Cereals	50391	52892	51675	51863	52154	52411
Oilseeds	7014	6554	7188	7304	7517	7800
Root crops	8149	5073	4410	4331	4252	4187
Gras and grazing	60410	55355	53532	52564	51368	50169
Other fodder	24191	20820	19509	18927	18107	17237
Other crops	17064	19125	19486	19433	19358	19301
Set aside and fallow land	13014	11315	12443	12642	12833	13008

Yields are forecasted to continue their modest growth over the projection period, with around +1 % p.a. for cereals and +0.8 % p.a. for oilseeds. Nevertheless, given the long projection horizon, average yields increase in the projection by +26% for cereals and +21% for oilseeds from 2001 to 2025.

Changes in the *beef sector* mainly depend on the interplay of the assumed continuation of the milk quota regime with projected milk yield increases which reduces calves availability and impacts on the forecasted beef cattle herd. Rather prominent increases in *pork* meat production in the seventies and eighties have already cooled down in the last decade, and EU production is projected to increase only slightly until 2025. In contrast to beef and pork, *poultry* meat demand and production continue in the projection their stronger increasing pattern. In both cases, relative increases are stronger in the new Member States. Forecasted demand and production of sheep and goat meat remain a small and decreasing part of EU meat markets.

Table 2: Reference run development for meat production

	1994	2001	2011	2015	2020	2025
Beef & veal	8861	8278	7579	7482	7351	7254
Pork	20041	21099	21993	22075	21916	21652
Sheep & goat	1169	1090	1061	1063	1049	1024
Poultry	8975	10658	11530	12039	12579	13111

The main focus of the outlook underlying this paper was the *state of the environment*. The projection results on activity levels have been used as inputs into other environmentally relevant models but they permitted at the same time to compute selected environmental indicators (nutrient surpluses for soils of nitrogen, potassium and phosophorus, gaseous emissions of ammonium, methane and nitrous oxide). These indicators are simplified versions of those calculated at a more disaggregate level in the CAPRI modeling system (for a recent exposition: Perez 2005).

Overall the changes are quite moderate because different impacts often cancel: In Table 2 we see for example that beef production declines but pork and poultry is expected to expand which tends to leaves most indicators stable, except for methane emissions which would decrease by some 5% in EU-23. In the crop sector yield increases also increase nutrient exports with harvested material but farmers increase mineral fertilizer supply at the same time. A remarkable improvement occurs in the surplus of phosphate which declines by about 20%, if past trends of improved management practices continue, mainly reflecting the assumption that more farmers account for organic phosphate when determining anorganic fertilizer application rates. A more detailed discussion of the environmental implications of the outlook is in Isoard, Witzke 2005.

7. Concluding remarks

The study underlying this paper resulted in a very detailed long run assessment regarding the state of agriculture and the environment which could be adressed only in very brief form above. The subtask to produce trend forecasts for such a long run horizon led to a quite sophisticated estimation methodology designed to minimise the risk of infeasible or otherwise unreasonable results which could not be presented in detail either.

The main focus in this paper was the idea to integrate a set of potentially contradictory external forecasts into a consistent reference run through the definition of a feasible parameter space and an objective function to steer the synthesis. This new methodology has a few advantages and disadvantages. In contrast to the original expert forecasts the CAPSIM simulation results are known to meet basic technological consistency constraints. Compared to the trend forecasts they will incorporate some economic reasoning because the shifters were not allowed to adjust freely to the set of external forecasts. The smaller the feasible parameter space the larger the role of behavioural functions in determining the final results. To put it differently: If a stand alone application of the model at hand would not appear recommendable for the questions at stake expert information and time varying parameters may be used to acknowledge issues left apart in explicit modelling.

The probabilities and additional weights in the objective function require and permit explicit statements on the importance of variables and on the confidence in the predictions of various sources. The compromise between different expert predictions is achieved through a set of equations rather than through an delphi process involving discussions among individuals intransparent to outsiders. Clearly, an equation based compromise will remain mechanical because human individuals may change their mind in the light of convincing arguments whereas supports are unaffected by a discrepancy to other forecasts. Even though solution time and transparency is certainly better than in a delphi process both the solution time and the model's complexity turned out demanding and require improvements. In spite of the last disadvantages the approach is promising in particular for smaller research groups lacking the time and financial possibilities for a lengthy process of evaluation, deliberation and discussion of forecasts by other agencies to end up at a synthesis by informal means.

8. References

Britz, W. (2003): Major enhancements of @2030 modelling system, Bonn: http://www.agp.uni-bonn.de/agpo/rsrch/at2030/@2030_2003.doc (25.11.2004).

Bruinsma, N. van (2003): World Agriculture: Towards 2015/2030. An FAO Perspective, Rome: FAO/Earthscan.

Conforti, P., Londero, P. (2001): AGLINK: The OECD Partial Equilibrium Model, INEA Working Paper No. 8, http://www.inea.it/opaue/wp8.pdf (11.12.2004).

CWFS/CPB, (Centre for World Food Studies / Central Planning Bureau) (2003): The CAP Reform Proposal of the Mid-Term Review: Decoupling with Strings Atttached, Chapter V, in: European Commission, DG Agriculture (eds.), Mid Term Review of the Common Agricultural Policy, July 2002 Proposals, Impact Analyses, Brussels: http://europa.eu.int/comm/agriculture/publi/reports/mtrimpact/rep_en.pdf (20.02.2003), pp. 141-169.

European Commission Directorate of Agriculture (2003): Reform of the Common Agricultural Policy - Medium-Term Prospects for Agricultural Markets and Income in the European Union 2003 - 2010, December 2003, Brussels: http://europa.eu.int/comm/agriculture/publi/caprep/prospects2003b/fullrep.pdf (25.03.2004).

European Commission Directorate of Agriculture (2004): Prospects for Agricultural Markets 2004-2011, Update for EU-25, Brussels: http://europa.eu.int/comm/agriculture/publi/caprep/prospects2004a/fullrep.pdf (29.09.2004).

FAPRI (2003): US and World Agricultural Outlook 2003, FAPRI Staff Report 1-03, Ames, Iowa: http://www.fapri.iastate.edu/Outlook2003/PageMker/OutlookPub2003.pdf (1.4.2004).

Golan, A., Judge, G.G., Miller, D. (1996): Maximum Entropy Econometrics: Robust Estimation with Limited Data, New York: Wiley.

Heckelei, T., Mittelhammer, R., Britz, W. (2005): A Bayesian Alternative to Generalized Cross Entropy Solutions for Underdetermined Models, Paper presented at the 89th EAAE Seminar, February 3-5, Parma, Italy.

Isoard, S., Witzke, H.P. (2005): Summary on selected results from an agricultural outlook for the '2005 EEA State of the Environment and Outlook Report', Working Paper, submitted to the XIth EAAE Congress in Copenhagen.

Lampe, M. von (1999): A Modelling Concept for the Long-Term Projection and Simulation of Agricultural World Market Developments - World Agricultural Trade Simulation Model, Dissertation, Aachen: Shaker.

Lucas, R.E. (1976): Econometric Policy Evaluation: A Critique, in: K. Brunner, A.H. Meltzer (eds.), The Phillips Curve and Labour Markets, pp. 19-46.

Münch, W. (2002): Effects of EU Enlargement to the Central European Countries on Agricultural Markets, Frankfurt a.M.: Peter Lang.

Pérez, I. (2005) Greenhouse Gases: Inventories, Abatement Costs and Markets for Emission Permits in European Agriculture: A Modelling Approach, PhD Thesis, Institute for Agricultural Policy, Bonn University, forthcoming.

Rosegrant, M.W., Meijer, S., Cline, S.A. (2001): International Model for Policy Analysis of Agricultural Commodities and Trade (IMPACT): Model Description, Washington D.C.: International Food Policy Research Institute (IFPRI).

Rosegrant, M.W., Paisner, M.S., Meijer, S., Witcover, J. (2001): Global Food Projections to 2020 - Emerging Trends and Alternative Futures, Washington D.C.: International Food Policy Research Institute (IFPRI).

Westhoff, P., Young, R. (2001): The Status of FAPRI's EU Modelling Effort, in: T. Heckelei, H.P. Witzke, W. Henrichsmeyer (eds.), Agricultural Sector Modelling and Policy Information Systems, Proceedings of the 65th EAAE Seminar in Bonn, March 29-31, 2000, pp. 256-63.

Witzke, H.P. (2005): Engineering Information in a Duality Based Agricultural Sector Model: CAPSIM, Poster Paper at the 89th EAAE Seminar, February 3-5, Parma, Italy.