Szent István University
Gödöllő

DOCTORAL SCHOOL FOR MANAGEMENT AND BUSINESS ADMINISTRATION

Thesis of (PhD) Dissertation

FORECASTS AND EFFICIENCY CALCULATIONS FOR AGRICULTURAL SECTOR MODELS

by:
Bunkóczi László

Gödöllő
2013

The doctoral school

Name: Doctoral School for Management and Business Administration

Discipline: Management and Business Administration

Head of School: Dr. habil Szűcs István
Professor, DSC of HAS
SZIE, Faculty of Economics and Social Sciences
Institute for Economics, Law and Methodology

Supervisor: Dr. Pitlik László
Associate professor
SZIE, Faculty of Economics and Social Sciences
Institute for Economics, Law and Methodology

………………………………………….	………………………………………….
Approved by Head	Approved by Supervisor

CONTENTS

1 INTRODUCTION

The paper (thesis) tries to show the arch of the work starting in 1999 from efficiency-calculations, through the agricultural sector models, and via the IACS[1] and led back to the forecasts and to the data that are accessible, and thus serve as a basis for research and forecasts.

Over the original aim (declared in the title) of writing this essay (efficiency-calculations and sector models), the following "phenomena" crossing the Millennium motivated. The phenomena are meant to be the processes in agriculture and formed situations.

The first three are stemmed from 2003, from Kapronczai's PhD[2] thesis-book (Kapronczai, 2003):

1. phenomena: (see in Thesis-book 2nd chapter):

   "At the end of the eighties, the reckless deregulation, when the existing data collection methods and databases were - hastily and carelessly – closed down, their lack later avenged itself.

   The mistaken belief that in the conditions of market economy, the demand for the amount of information is reduced. (...)

   Last but not least, that the agricultural sector has become "more complex" to display it by the information systems. This on the one hand - I mean - that while in the past 3-4 thousand farms existed, monitoring them to get the economic picture for the agricultural economy as a whole, today almost ten times much should be observed. On the other hand, the product chain connections were more controlled and follow able, which also made it easier for showing the real processes."

The deregulation process outlined (e.g. data collection is stopped) in prosperity can be positively judged. The negative perception in general will begin to come to the forefront, when the deregulation due to market turmoil/addition (e.g., overproduction, lack of external import, export options for shrinkage, etc.) are growing up and more grievances (even worldwide) is added by the crisis. Since the autumn of 2008 on the basis of the general knowledge has been crisis, and also in the second quarter of 2012, for example, increased acute crisis (e.g. Hungary, 0.2% in the first quarter, and a decrease of 0.7% of GDP[3] on quarter/quarter) reported the Central Statistical Office (CSO[4], 2012). Therefore, the risk of negative scenarios (which significantly exceeds the potential benefits) can be significantly minimized by reversing the deregulation process (e.g. with the introduction of system-level modeling in the agricultural sector, and with regulations).

2. phenomena: (see in Thesis-book 5th chapter):

   "Along the scientific examination of the financial information systems it was established that this is the area where the available data is the most controversial. The National Tax Office's database hides numerous uncertainties. This arises from the unreliability of the basic data. The firms bounded for tax declarations, tax returns have numerous opportunities "to conceal" their revenue in order to avoid or minimize tax paying, which causes the creation of a database reflecting the true financial processes behind hardly."

The APEH[5] (Tax Office) /NAV[6] database itself hides no uncertainty, but consists the unreconstructability. In fact, the main method for taxation is the self-declaration of tax, which has to be performed per tax per period. This may cover from the free of tax (e.g.: free of declaration), across the annual single declaration (e.g. for primary producers if the turnover is less than 600 thousand HUF than subjective VAT exemption) ranges till the monthly reporting commitment (e.g. VAT, taxes on labor and wages). Obviously, the NAV tax declarations are not organized on area basis, so if 500 ha

[1] IACS Integrated Administration and Control System, the whole of the Single Area Payment Scheme's (SAPS) institutional, legal and other frame system, starting in 2004

[2] PhD: *„philosophiæ doctor"*, the only gain able scientific degree in Hungary in nowadays

[3] GDP: Gross Domestic Product, all the products and services being produced and given within the borders of the country, usually in one year

[4] CSO: Central Statistical Office

[5] APEH = Tax and Financial Controlling Office

[6] NAV: National Tax and Customs Office

land is cultivated, it is conceivable that only 500 thousand HUF annual net output value invoiced. Till the NAV cannot refute it factually, the suspicion generation (e.g. fraud detection) potential remains low. However, with introducing the sector models and their farm / product settlements (column and row) the potential could be significantly increased – on its self-controlling declaration basis.

Another problem layer is that for information providing towards the CSO's not all are bound in agriculture, thus there is no theoretically chance to compare data provided to the CSO and to NAV. For detecting inconsistencies (see MIMIR study, 1998: Consistency committee) institutional and regulatory frameworks are missing, but about methodological deficiencies we can't speak about since the existence of agricultural sector models.

3. phenomena: (an article which is based on ARI[7]'s fresh study from 2011):

   "The tax system for the farmers do not encourage them to evince neither income nor the expenses, beside this a minimum social security payment - writes the Economic Daily. This is evidenced by the fact that as the result of the tightened wealth accumulation investigations since 2007 the self-employed enterprises in their declarations doubled their income, while the revenue grew only 12 percent from one year to the next. The lack of social security payments lead to serious long-term social tensions. The study found that farmers receive a significant amount of 40-60 billion HUF in assistance through the tax system which equals the one fourth-one third part of the subsidy that they receive as agricultural and rural development, and other support" (index, 2011a)

In practice, all can be traced from the previous (2nd) phenomenon. Much of the aid is guaranteed EU[8] funding, it's third / fourth is concerned to be 40-60 billion HUF. Comparing to a reference, to the governmental interest expense in 2012 is expected to take 1.049 billion HUF (i.e.: 7%) (gazdasagiradio, 2012) out of the total budget, which is 14.899,8 billion (100%) HUF, agreed as the 2012 year budget law on the 28th November in 2011. That is (not specifically studied in this thesis, but) a worth to study basic economic situation, based on the "many little makes a mickle" concept fits in connection to concentrate for any billion in budget planning / monitoring.

4. phenomena: in 2011, according to the news published in abridged the GDP contribution of the agriculture was 3%, which makes the double of EU average (Agroinform, 2011). After our general knowledge, this value (e.g.: in the eighties) was more than 10%. Quoting further: "the entire agricultural sector - including farmer servers, producing raw materials for the agricultural industry and production assets of the host processing industries - makes up the 15% of the domestic economy" (Agroinform, 2012).

The factual perception of the 3% share, as much or little, probably mostly a matter of perspective. After the general historical knowledge, it is said to be few, and after the EU average, too much.

5. phenomena: The following quote is derived from an online newspaper's article (index, 2011b) from 2011, entitled: "A third of cereal sold by VAT fraudsters'

   Experts say, the technique is becoming more and more prevalent - and also the most difficult to grasp - that the fraud and income hiding starts at the producers. Some of the grain already denied and later sold mostly for cash through the well-established only for few a month long existing trader firms. These firms sell the goods invoiced but the payment of VAT tax fails to have. The tax authority usually starts the control by the greater players at the trading chain end, but when they get to the front of the chain to the phantom firms, by that time, these companies have long finished their activities. (index, 2011b)

As a result, on 1st of July 2012 the NAV introduced the reverse taxation of VAT for cereals, which actually suppressed this kind of committing method, however the root cause, which is the excessively high VAT tax rate for basic food in the EU, upheld it, and even increased it by 2% which strengthens more the grey and black economy.

---

[7] ARI: Agricultural Research Institute

[8] EU: European Union

6. phenomena: Under the auspices of the EU with 27 countries, the current CAP is driven by diverse interests. The current "we don't rule" (Székely, 2011) state doesn't lead for stability for countries of South-Europe and for us, which is amplified by the roller coaster of the weather influenced quantity produced and the changes of the sown area.

In the 27 countries the current valid CAP is not unified, there's everywhere some national playing field, which only cause that the rich(er), normally industrialized countries can protect/subsidize appropriately their agriculture, while the less wealthy, can't or do not make the same. Thus, the different attitudes of French, a German, a Spanish, an Italian and a Hungarian farmer within one sector. This, than the periodic reforms of the CAP and the Commission's decisions to incorporate into any longer term simulation is a complex challenge, as complex as the good and responsible planning. The latter would be the basis for short-and long-term effective production. This paper contributes to this latter idea that the annual change of yields, prices and area are predictable at higher utilizable accuracy levels than known up till now, on the base of ex-post calculations. The study area is within the agriculture, the (arable) crops, which are the most affected by the weather.

All PhD thesis aims to achieve new or recent scientific results. There are new and more recent scientific findings in vain, if:

- on one hand they are calculated on the basis of incorrect (incomplete, invalid, etc) data, and get published,
- secondly, the reality is far more complex than the models can be created,
- thirdly, the reality is distorted at certain points, which in some cases because of its uniqueness can't be modeled.

Based on Kapronczai (2003), earlier 3-4 thousand farming units were, while now there are about 87 thousand units (http://miau.gau.hu/fadn field E = all filter = all year = 2009) to monitor, so it's follow up assumes multiple workload. In the absence of reliable data, it is simply impossible, and if we had reliable data, include into a cellular automation (as an emergent system) to recreate what all players make (for what and why), also falls into the impossible category.

The major part of the theme, the economic models, including the agricultural sector models (sectors, yields, inputs and their level, input and output prices, accounting these a closed system) deal with the same problem at national or even at EU level.

In reviewing this topic, as well as reviewing the "development" of the models, it can be well seen that in case of dynamic mapping nobody can be sure whether:

- actually the reality was modeled,
- the internal logic is actually correct,
- whether the potential derived state is actually ideal state (equilibrium) or not.

The "agricultural sector-models" (shortly as ASM 's) has the advantage that they are able to describe almost the whole agricultural sector for countries with the data stored inside. So here one can see almost all the data of important sectors. Planning methods (e.g.: Work-Table Programming hereinafter referred to as MTP) is useable here, but of course depending on the resolution of regional, national or EU-27 size. The goal is here that with appropriate internal model (the model itself is the soul – the programming algorithm) would be pretty much easier to model the intricate and complex reality for +1, +2 and +3 year(s) if all affecting factors are present. The real problem with these models is that the returned values by the model such as "endogenous variables", largely depending on the inputs fed into the model as "exogenous variables"- and will be calculated on the basis of these exogenous variables, which are mostly linear trends and/or expert opinions bases. No one has ever validated the values of these exogenous variables, and model validation does not mean which anyone may think for the first round.

According to the literature, the validation criteria is not clear, so it is unclear whether for the results of a certain model run, the "experts" would say the same, and would think the same too about it.

For us, the ASM 's primary lesson of the inherent data fortune and the data structure itself.

At that time, when production-, input-, output- and price data will be available at SPS or SAPS supported table/plot level (e.g. MEPÁR browser, under http://www.mepar.hu), will come that step to fine tune the model's internal algorithms and include the appropriate forecast methods described in this paper. Including the forecasts instead the up till now the mostly used trends, the "exogenous variables" shall be actually good (direction hit and close value) and consequently, the model shall give actually good values back as forecast, as simulation run or as scenario run.

This paper's declared aim to demonstrate, that it is possible to forecast the sectors "corner stones" (yield, price, area) values with relatively acceptable good predictions, against the previously widely used linear trend, which in the tests a total of 9.73% was shown to be superior to others. In other words, the primary objective is restoring the fundamentals, the inbuilt systems and allocation algorithms of the models should not be tampered.

The well foundation of the predictions is essential in future planning and for design methods. It may be that since a time ceased to be an explicit claim for such a thing, but in competition and in an escalating situation (examining the EU and worldwide) on the basis of a credible vision a "right" decision may contain a significant advantage for any farmer/entrepreneur. This may leave time to prepare for those cases, when no particular regulation exists and everyone waits the state for solution/help.

Beyond the earlier idea, the DEA (Data Envelopment Analysis) efficiency calculation can be related to the topic, that by default, purely technical efficiency are calculated, where the input factors prices as a monetary determinants (e.g.: inflationary pressures in the nineties in Hungary) does not matter. This has the significance, that in case of a stable exchange rate regime, each generating unit's actual efficiency (in DEA terminology: DMU - Decision Making Unit) is well-calculable, resulting in profitability significant differences in accordance with e.g.: the size in any manufacturing sector. In long term for all countries and for the EU too is important to use resources in the most efficient manner as it's possible. The DEA-based efficiency calculation has a yield, that with weighing the inputs quasi production functions are created, which can be utilized in planning, but similar to the forecasts in this case the missing future (DEA) efficiency is what needed to know in advance.

The efficiency calculation importance it is also seen that during building agricultural sector models, based on McCarl (1982) suggestion Jonasson and Appland (1997) begin to group the farms in which one aspect is the efficiency of operation, because the "overall average farm" in reality does not exist. This idea shows just how complex this issue is and that the issue is from the basic statistical data collection in the agricultural sector for models and covers almost everything between the two end points (accounting, production function, efficiency, forecasting).

Identified problems and task assignment:

1. In sector-modeling arises the problem not actually mean/average "farm", which is true for almost every production unit. From the agricultural sector-model we get to a single production unit, in aggregate (country, region) modeling we need to face the fact that an entire country / region assuming one unit, the overall level of abstraction is excessive high. Thus the observable production units / representative sample group, shall be grouped which can be made on the basis of efficiency, but in this is case it's realistic only by sectors. The real solution would be monitoring each production unit, but it is not feasible because of the current known production unit of 87 thousand. Because of the well-known and commonly used DEA method is time-consuming, basically unfit for any such use, that's why the primary goal is to replace it.

2. Due to the inadequacy of current data collection and control a data collection methodology would be required for collecting production, yield, input, output and price data for the whole sector and data for employees as well. A basic criterion is none other than the proposed methodology should be adaptable to any size of enterprise or site.

3. For the agricultural sector modeling and for planning the problem of forecasting the exogenous variables (corner numbers: price, yield, area) and the problem of validating their values. Instead the current general ("quasi") linear trend based forecasting, better alternatives have to be shown.

4. All of the forecasted value has to be validated. This process should be built into the forecasting methodology. The steps of validation must be defined and have to made them applicable.

5. These in fact better alternatives that are able to deliver data to widely known planning methods as "exogenous" data and from these data gets these methods capable for the production of surplus value.

Thus, the thesis is not an investigative reporter routine collection, but some of the methodological innovations of the establishment of a consistent plan/fact data fortune!

2 MATERIAL AND METHODOLOGY

2.1 Material

The DEA analysis was performed on the SPEL database, from its 1999 state, the complete data capture was made from its 2000 state. The data is stored in so called "tab" files in which the last instance of approx. took 60 MB of space. To extract such data from SPEL, the Daout.exe had to be used which had a pivot styled selector surface and ran only under Windows 95, so by changing the commonly used operating systems (Win95 and Win98, Win2000 and WinXP, Windows Vista, Windows 7, Windows 8) the recovered 2000 complete putative database remained.

The forecasts were made on the time series of FAOSTAT database website, exactly under http://faostat.fao.org, where production statistical data can be found. Caused some concern that the yield was original in kg/ha than later in gram/ha, and most recently in hectogram/ha, and the previously published data has changed somewhat, so all the forecasts are made on the data(base) downloaded in 2008. So they are comparable without any restrictions, the time series are identical.

Other features of the FAOSTAT, that price, unlike the other 2 data type (yield, harvested area) were available only from 1991 to 2006 while the others from 1990 until 2007, and further specificity is the case of Germany where all time series are shorter with one member.

The work was made under the 2000-2003 version of MS Excel spreadsheet software, where it's possibilities was mostly tuckered, with largely set programming, with solver runs and with programmed solver runs.

2.2 Methodology

2.2.1 The mathematical, theoretical background of DEA analysis

The idea of DEA comes from FARREL (1957) and reforming it as mathematical programming problem to CHARNES, COOPER and RHODES (1978). Given a certain number of production units, called as Decision-Making Unit (DMU's). The DEA method determines the border of efficiency from the efficiently working unit's example. The marginal efficiency reflects the operation of existing plants. Those units which operate on marginal efficiency are not considered to be effective. Measuring the effectiveness of the units is done by dividing the sum of the weighted outputs with the sum of the weighted inputs, and then finding the maximum of this quotient by changing the weights. Due to the constraint this rate rate of each unit is less or equal to 1. The DEA model for each production unit occurs as a non-linear fractional programming problem. The optimization task performed in the non-linear fractional program is the following:

(1) $$\max_{h_k} = \frac{\sum_{r=1}^{s} u_r y_{rk}}{\sum_{i=1}^{m} t_i x_{ik}},$$

concerns for:

(a) $$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} t_i x_{ij}} \leq 1, \text{-re}$$

(b) $u_r, t_i \geq 0$

(2) $$\max_{u_r t_i} \sum_{r=1}^{s} u_r y_{rk} - \sum_{i=1}^{m} x_{ik} t_i$$

concerns for:

(a) $$\sum_{r=1}^{s} u_r y_{rk} - \sum_{i=1}^{m} x_{ik} t_i \leq 0,$$

(b) $u_r, t_i \geq 1$

Where:
h_k = relative efficiency of the k DMU
u_r = the weight of y_r output, $u_r \geq 0$
t_i = the weighs of the x_i inputs, $t_i \geq 0$
y = the output of one production unit (DMU), $y \geq 0$
x = the inputs of one production unit (DMU), $x \geq 0$
j = the indices of the production units (DMU), j = 1, ... , n (n = number of DMU's)
i = the indices of inputs, i = 1, ... , m (m = number of inputs)
r = the indices of outputs, r = 1, ..., s (s = number of outputs)
k =(special) standalone production unit (DMU= Decision Making Unit)

The (2) type of equation is called a multiplier formula within the programming problem. Using the principle of duality, we get the equivalent envelopment-formula, which has less restrictive conditions to take into account, so it is easier to solve. The linear programming problem's dual formula is equation (3):

(3) $\min_{\theta_k,\lambda} \theta_k$

concerns for : (a) $\sum_{j=1}^{n} y_{rj}\lambda j \geq y_{rk}$,

(b) $x_{jk}\,\theta_k - \sum_{j=1}^{n} x_{ij}\;\lambda j \geq 0$,

(c) $\lambda_j \geq 0$,

Where,:

θ_k=Debreu-Farell-efficiency value

λ_j = constant numbered weight vector

Equation (3) shows that at the featured "k" production unit the minimum value of input efficiency is θ_k is determined by the model. θ_k shows the "k" production unit's Debreu-Farell efficiency value, which should satisfy even the $0 \leq \theta_k \leq 1$ criteria also. The weighted combination of the output neither for a single r output shall fall below the level of the total input of the "k" production unit. In addition, a single case of a weighted combination of "i" input shall not exceed the total level of the inputs of production unit "k".

2.2.2 Simulation of DEA

Since the classical DEA method is quite cumbersome and originally running was only possible in local environment under GAMS, so DEA simulation was worked out in 2001. Due to its simplicity it allows to be operated for/via the web. The results obtained on the basis of Pearson's correlation coefficient (0.88) shows a strong correlation between the calculated values and classical DEA values.

The stability of the solution is showed by the case of Austria and Denmark, the only naturalistic quantities of N, P, K's weights were set to 1, while the monetary influenced "variable energy costs" was set to 0,001, and the correlation between the original DEA values and the simulated DEA values (relative between [0;1]) is 0.97.

The initial equation is used for the method. Basically an output existence is assumed and therefore no weight is needed for the output value. Moreover, the method assigns for all DMU's inputs only one weight per input.

1. For one DMU: $\sum x_i{*}t_i = y$, that is, the sum of the weight multiplied inputs (sum-product) equals with the output. Since both theoretically and practically a most effective DMU exists, and in case of well-adjusted weights it's true only there. The function is basically a production function.
2. The efficiency per DMU is: $\eta_j = y_j / \sum x_{ji}{*}t_i$ where each $\eta_j <= 1{,}00$. In case of a normal dispersed case collection this is true only once. All other cases should be less than one, since in other cases the input is not used efficiently, more output could generated, thus the use of inputs is "wasteful".
3. Solver's solution where the maximum amount of efficiency values are calculated, with a constraint that none of the efficiency values shall exceed 1.

$$Max : \sum_{j=1}^{n} \eta_j = \sum_{j=1}^{n} \frac{y_j}{\sum_{i=1}^{m} x_{j,i} \cdot t_i}$$

Condition: $y_j = \leq \sum_{i=1}^{m} x_{j,i} \cdot t_i$ for all j=1, …, n

Where,

- n – number of DMU-s
- m – number of inputs (m_j – the number of the j^{th}. DMU's inputs)
- j index - index of DMU's (j = 1, … ,n)
- i index – numbers of inputs for one DMU- (i = 1, … ,m or m_j)
- $t_{i,j}$ – the weight for the j^{th} DMU's "i" inputs $t_{i,j} \geq 0$

The DEA simulation run experience was, by calculating the efficiency not by objects, (such that being the "peer"), therefore efficiency values calculated in the DEA simulation are more stringent, the method is less "permissive".

2.2.3 Expanding SPEL data, and programming settlements, reckonings

In the summer of 2000, the complete SPEL database was collected down into structured form:

- 35 main plant products, for 21 countries, + EU11 and EU15 s averages, and all from 1973 until 1998, 26 years
- 11 main animal products, for 21 countries + EU11 and EU15-s averages, and all from 1973 until 1998, 26 years.

A complete database of the ECU/plant, ECU/animal, national currency/plant and national currency/animal (4 files) can be found.48 Mbytes length. The querying and storing into structured work meant several weeks, but programming the settlements itself minimum took 1,5 but better 2 months (summer of 2000) time/work.

Programming the settlements/reckonings resulted 4 macros (1 row and 1 column oriented for the plants and animals in national currency and in ECU depending the adjustment) which have source code about 240 thousand characters (in GTK diploma length approx. 3 diplomas). Along the development these settlements had a positive echo by the German part, that the SPEL data could be seen in this form. Since then, it has become clear that in this form is not likely that the developers could see the data as they only received it from the partner countries, it was integrated (loaded) into the database, and the developers further developed their program codes.

According to the literature part about SPEL, the four quadrant matrix (product formation, expense formation, expenses use and the use of products), column and row oriented reckonings/settlements were programmed. The programming was based on the SPEL methodical manual (SPEL system, Methodological documentation (Rev. 1), Vol. 1: Basics, BS, SFSS, WOLF W. (1995)).

2.2.4 Consistency

Quote from a report:

> *'Task was under the fall of 2002 to the department (GINT) to make forecasts (statistics in the absence of previous partial data) for 2001 for certain products, which can be expressed beyond the retail price of certain costs, respectively the evolution of price components in Germany. For this the SPEL data was used, where the production cost per components time series (natural and price levels) were available, respectively, and prices and yields. The forecast was based on the function of the spreadsheet software "trend". The degree of self-control was the limited difference between the forecasted components and the resultants (sum of the components). The experience states that the individual cost component's trends like price and volume trends showed probability of relatively low, respectively changes so the results of the bi-directional prediction (in a deteriorating way) is almost the same for 1-2-3 years." (Pitlik, 2002)*

During the work linear trends were used, and the results were consistent with the direction of the primary aspects of the 2001 forecasts, both in case of inputs used or their price, or in case of outputs and their price.

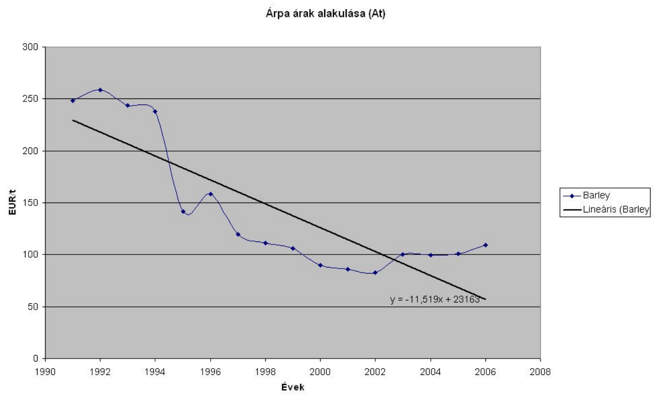
Since data for 2001 was available already in 2002 for Germany, that were clues and checkpoints, so it can be quite accurately expressed, that predicted missing values were in harmony to reality in all respects.

2.2.5 Forecasts

The following methods were used during the tests on the same testing data:

- Linear trend
- [Polynomial trend] – was rejected, as the CAPRI also rejects it, although they stopped at second degree
- Chronological weighted trend – rejected also, because of the intensive polynomial effect
- Wave-functional trend (later may be found as: „sine" also)
- Similarity analysis based forecasts (COCO)
- Fundo-chartist approach (in 3 stages/steps)

2.2.5.1 Linear trend



1. Diagram Linear trend fitting for the whole known period (own representation)

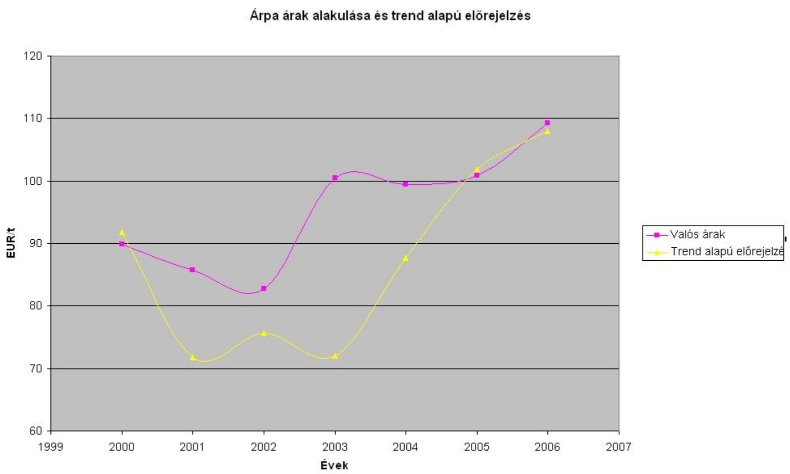
In general, fitting the well-known y=mx+b linear function for the known time-series. It's main attribute is the "m" slope (with "+" means growing, with "-" means decreasing, by 0 it means constant height and independent from x line and) and "b" is a constant that specifies at which height the "y" axis which intersected by x = 0.

The fit is most accurate when the sum/amount of the distances of actual and fitted (discrete) data points (without sign or squared distances) the minimum is. Because of the trend function of Excel solves the fitting problem, in particularly not needed to deal with parameters. As referred in SPSS description, fitting onto time series, splitting it into two sections, such as learning and testing are irrelevant because it can be made only in one manner. Then in the test period, we get something.

A graphical example for the application:

On the above graphical example of fitting the trend was made for the whole period, where trend is below than above and finally below the real values again. Within market conditions, the key value of measure the prediction is the year/year comparison of direction hit (fail or hit the growth or decline). Easy to see that the trend values would have been quite misleading, but calculating back values from the downward trend to the last known, then 12 times out of 16 (75%) result could have been reported.

Comparing the real price evolution trends and trend changes (the change of the real is the same as the trend change) only 40.82% result was obtained. Written differently: where the two functions are moving in the same direction between two years, then direction hit is found, in other case there's no hit.

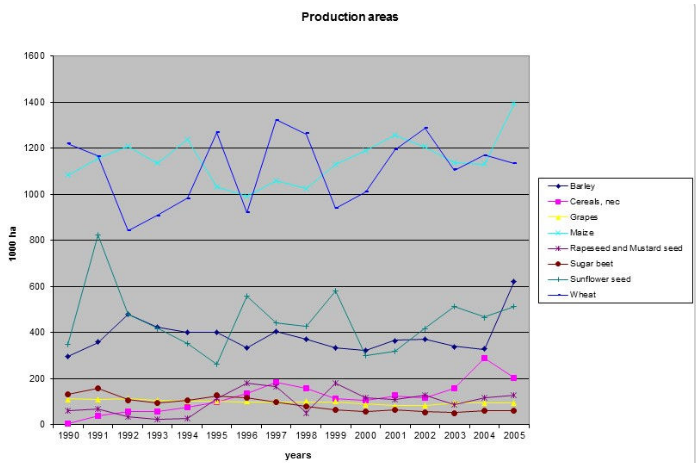


2. Diagram Linear trend fitting for known data of 5 years, the 6th is plotted (own representation)

Beyond the direction hit, the differences and the adjusted differences are also very important. These differences significantly affect any kind of optimization model's the result, due to a significant price underestimate the calculated Gross Margin will be much lower, so the model necessarily underweights the sector, so next year profit loss will occur (see MTP programming). In inverse case above-weighting which cause losses later.

In CAPRI also fixed that they stay by the linear trend-based projections, as forecasting for 6-8 years the polynomials have a very negative effect: they keep to the plus or minus infinity. The problem quickly becomes quite sharp. Especially if it's tried like in CAPRI. From three years average ahead 8 year (2003-2005 → 2011).

2.2.5.2 Wave function fitting



3. Diagram Yearly changes in growing areas (source: http://faostat.fao.org)

The idea of using trigonometric functions goes back to the beginning of the 90's (Pitlik, Dissertation, 1993) and to the autumn of 2003, when it was demonstrated that for 5-6 year time span a combined sine function had very good fit. After the minimum of the 2003 year soft-wheat yield (2662 kg/ha), which was well-followed by the function, the function calculated more than 5 t/ha yield ("predicted in advance") for 2004. The FAO data show that (according to the 2007 data query) in 2004 in Hungary the mean yield was 5139.9 kg/ha.

On the 3rd diagram (above) with a bit of abstraction, can be viewed through the "largely" normal fluctuation of major crops growing areas of time series and also that they move in a well-definable range of values.

The next equation the solution for time-series with trend:

$f(t)=\sin((t-p_1)/p_2)*c_1+c_2+c_3*(t-t_0)$ where:

t: the value of a given year

p1: parameter for shifting the zero moment of the period

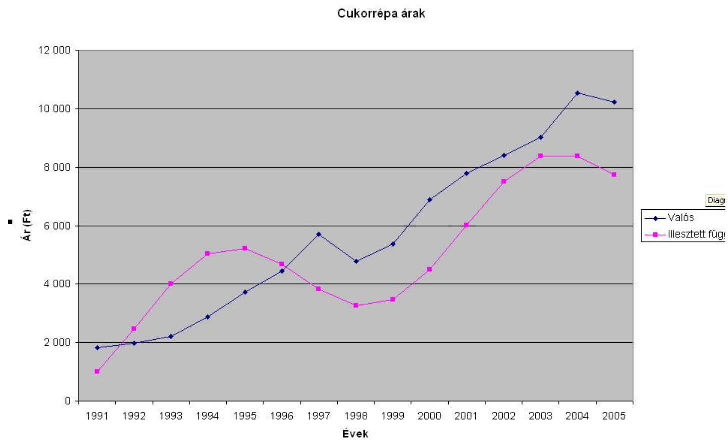
p2: expanding or reducing the period parameter

c1: the amplitude of the wave, the known values of the standard deviation of time intervals

c2: initial height of the baseline, average of the first 3-4 known values

c3: the calculated slope of the difference of the first and the last 3 values of the known time interval

t_0: the first known year



4. Diagram Changes of the price of sugar beet and wave function fitting (data source: http://faostat.fao.org, own representation)

Finding the parameter combination which approaches the known interval closest theoretically is possible with the built-in Excel solver, but in fact, the solver cannot solve these problems, as it often stops due to inappropriate target change, so a different solution had to be found.

The systematic search became the useful practice, which moves the two parameters independently in proper intervals with certain steps. So 101 * 10 parameter combination's all data per sector (plant, p1, p2, difference, learning%, test%, the proportion of test/study) is recorded by the macro. In case of difference, the sum difference between the real values and the fitted function's values is meant in the learning phase. A typical figure thus shows (see 4th Diagram).

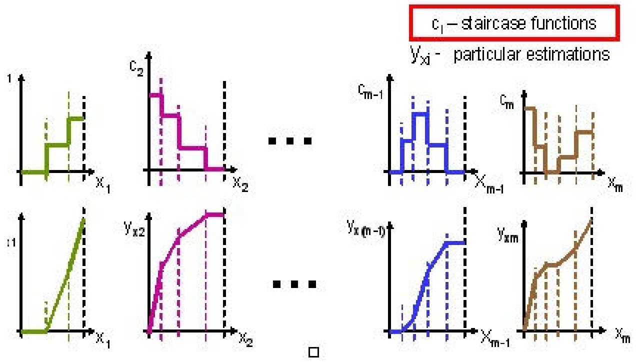
Thus, during the run for search of parameters for the price, good learning and similar good test values are obtained as in the first run, and all the differences in known time interval between the actual and predicted values decreased significantly.

As this quantity of time series and results cannot be considered as a reference value and the results were better than expected, so the data to be listed, was the results of a 158 set of time series (52 harvested area, 52 yield, 54 price, with 16-17-18 elements) tests performed with all methods on EU countries as well as for Switzerland on the most decisive 5-6 plants per country.

The practical solution for the wave function trend fitting, is made that way, that having a pretty good but not exceptionally good result for the known stage, and from the list looking for a relatively low deviation value and selecting it, got a "test" result value. Compared with other methods of determining learning/test results was more subjective, but not exclusively 90-100% of study or test values were selected. The goal was rather, to model the reality.

2.2.5.3 Forecasts based on Similarity Analysis (COCO method)

$$y = c_1(x_1)\cdot x_1 + c_2(x_2)\cdot x_2 + \ldots + c_{m-1}(x_{m-1})\cdot x_{m-1} + c_m(x_m)\cdot x_m$$



1. Figure Types of staircase functions (Bánkuti, 2010)

The mathematical interpretation of the method was given by Györgyi Bánkuti on the summer of 2010, so only its main elements will be published:

The novelty of the method lies in the fact that beside changes (increasing trend, decreasing trend and stagnation) previously describable only by general linear regression, more complex changes are describable with. In case of 5-10 element long time series is possible that the changes we are seeing is really more complex than just growth, decline or stagnation.

The following diagram (Figure 2) presents the solvable basic linear programming task of the similarity analysis.

Over the mathematical interpretation the method through ordering time series into a matrix (e.g.: 5 vectors, which last elements always x_{t-5}, x_{t-4}, ..., x_{t-1}, and the rest of the vectors are always the prior elements) seeks the closest time series downwards, to minimize the differences between the known and calculated values through a gradual "value-splitting" by each vector and returns an output value for each vector. These values are usually summed (additive process) to give a realistic value for the unknown "x". Then this "x" can be compared to the real one.

$$\begin{bmatrix} \underline{\underline{E}}^{-+}_{(n-1)xn} & \underline{\underline{0}} & \underline{\underline{0}} & \underline{\underline{0}} & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{E}}^{-+}_{(n-1)xn} & \underline{\underline{0}} & \underline{\underline{0}} & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{0}} & \ddots & \ddots & \vdots \\ \underline{\underline{0}} & \underline{\underline{0}} & \cdots & \cdots & \underline{\underline{E}}^{-+}_{(n-1)xn} \end{bmatrix} \cdot \begin{bmatrix} \underline{s}_1 \\ \underline{s}_2 \\ \vdots \\ \vdots \\ \underline{s}_m \end{bmatrix} = \begin{bmatrix} \underline{0} \\ \underline{0} \\ \vdots \\ \vdots \\ \underline{0} \end{bmatrix}$$

$$\begin{bmatrix} \underline{\underline{Y}}^1_{nxn} & \underline{\underline{Y}}^2_{nxn} & \cdots & \cdots & \underline{\underline{Y}}^m_{nxn} \end{bmatrix} \cdot \begin{bmatrix} \underline{s}_1 \\ \underline{s}_2 \\ \vdots \\ \vdots \\ \underline{s}_m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_m \end{bmatrix}$$

$$Z_{lin} = \underline{1}^T \cdot \underline{s}_1 + \underline{1}^T \cdot \underline{s}_2 + \ldots + \underline{1}^T \cdot \underline{s}_{m-1} + \underline{1}^T \cdot \underline{s}_m - \underline{1}^T \cdot \underline{y} = \min.$$

$$Z_{Abs} = \sum_{i=1}^{n}\sum_{j=1}^{m} Abs(s_{ij}) - \underline{1}^T \cdot \underline{y} = \min.$$

$$Z_{Squared} = \sum_{i=1}^{n}\left(\sum_{j=1}^{m} Abs(s_{ij}) - y_i\right)^2 = \min.$$

2. Figure General linear programming model for the similarity analysis (COCO) (Bánkuti, 2010)

In earlier Best Student's papers and thesis's up to 85-93% direction hit were obtained, which certainly noteworthy and thus the method itself testable.

The solution was made by programmed Solver runs, so the Solver ran about 158*6 times as the time series test phase is mostly 6 years long. The data was placed at several worksheets. The first page, where the longest time series were placed, the Solver gone for 70-80 steps, while on the last page after 10-20 steps stopped with solution.

2.2.5.4 Fundo-chartist approach

This method is designed for expanding the chronological weighted trends. Here, the chronological weighted values of all elements (plant) specific to the plant is divided by a parameter value (plant effect for the others, inverse effect), the values are summed and then multiplied by a parameter that is only specific to the plant (scale setting), and all this is made by countries. The difference between the first and the second parameter is that while the first (five or six - depending on the number of plants) affects all of the plants, while the other only have effect for one specific plant's certain attribute (yield, price, area) and plays an important role in, that the values should be neither too small nor too large.

$$y_{i,t+1} = \sum_{i=1-n} [(y_{i,t-4} * s_{i,1} + y_{i,t-3} * s_{i,2} + \ldots + y_{i,0} * s_{i,5})/p_{1,i}]*p_{2,i}$$

Where: $y_{i,t+1}$: t+1 year reported value of the i^{th} plant (yield, area, price),

s: used weights

p1 and p2: parameters

Compared to the weighted chronological trends, it turns up or downwards much consolidated out, so that kind of consistency/plausibility problems shall not be overcame, as earlier. A multivariate approach is advantageous in that respect, since all the previous data of all plants to fit on, has a strong limitation/constraint effect as well.

From this three types of solution were made. In the first version the fitting was made once for the learning period for all the plants (5-6 plants per country) and with the used weighting the test was made for 6 years in advance for 5-6 plants. In this case, actually we get a plant combination ahead for 6 year, with yield, area and price for each country.

In the second case, the fit was per plant, using the entire learning phase database; one is made per plant for a 6 year test period as ahead projection.

At the third case there is fit per plant and per year (annual), so in a country where six plants are presented up to 6 years, there's 36 times running of the solver to more precisely learn the time-series and project out to one year in advance. This solution's values were included in the results.

3 RESULTS

3.1 DEA simulation

During the simulation of the original DEA efficiency calculations, managed to simulate it with a correlation coefficient of 0,88 for the technical effectiveness of the total EU15 soft wheat production database. In addition, the original database consist confused/corrupt data, which can be handled better by the classical DEA than the simulation, as in the original DEA only one value and run is influenced by seriously, while in the DEA simulation, because of the one run – all value. Taking this into account the 0.88 correlation value is better than expected.

1. table The correlation of DEA and DEA simulation (own calculation)

Country	Year	DEA efficiency	OUTPUT1 (SWHESWHE)	INPUT1 (NITF+NITM)	INPUT2 (PHOF + PHOM)	INPUT3 (POTF +	INPUT4 (CAOF)	INPUT5 (ENEV)		
				22,02	10,92	16,88	27,81	0,00		
Svédország	1981	0,39	3 226	148	71	116	104	23 865		
Svédország	1982	0,52	4 446	154	74	117	106	29 032		
Svédország	1983	0,59	5 239	160	77	121	109	31 046		
Svédország	1984	0,59	5 210	167	80	117	111	33 852		
Svédország	1985	0,45	4 043	173	83	115	113	37 884		
Svédország	1986	0,6	5 373	180	86	113	115	35 280		
Svédország	1987	0,51	4 637	187	89	117	117	42 735		
Svédország	1988	0,53	5 000	194	93	122	119	41 393		
Svédország	1989	0,63	6 014	201	96	122	122	45 400		
Svédország	1990	0,65	6 449	211	101	129	124	51 000		
Svédország	1991	0,61	5 859	216	103	121	126	52 234		
Svédország	1992	0,54	5 325	223	97	122	128	51 940		
Svédország	1993	0,57	5 891	231	106	133	130	56 644		
Svédország	1994	0,5	5 342	239	114	143	132	56 948		
Svédország	1995	0,5	5 945	246	118	167	135	61 200		
Svédország	1996	0,5	6 191	254	121	181	137	75 768		
Svédország	1997	0,47	5 976	262	126	196	139	78 716	Correlation:	88,16%
Svédország	1998	0,48	5 906	271	129	175	141	80 736	Pearson:	88,16%

With fixed weights valued 1 (except for the monetary effected ENEV of energy, where this is 0.0001), by Denmark and Austria 0.98 correlation values are gotten between the DEA and the simulated DEA.

The relative values are between 0 and 1 and these are the simulated values of the DEA efficiency. Without simulation run so stable solutions are available on this basis.

Necessary for offering onto web move the simulation itself, as the classical DEA execution time required hours on the same database, while in the simulation solution by Excel solver, only seconds. The (potential) users are not willing to wait for minutes or hours for the result.

3.2 SPEL database, method for settlement and scheme

The SPEL column and row oriented accounts scheme is none other than the logic inside SPEL. We need to be able to account with all produced or existing products and with all inputs from creating till use, in natural and in monetary way too.

The data can be considered as a table, most transparent form which can be divided into four quadrants:

- upper left quadrant - product formation
- top right quadrant - product use
- lower left quadrant – expenditures/input consumption
- right lower quadrant – expense/input formation

The **table rows** above contain the products listed in the agricultural category (e.g.: wheat, corn, milk, eggs, etc.), respectively, in the lower part of the expenses used to produce them (NPK, seed, feed, chemicals, energy, services, etc.).

The columns of the table on the left represent each sector respectively, on the right the use of them or the place of formation (own consumption, market sales, or market purchasing, processing, stock changes, etc.).

Obviously the **four quarters (naturalistic data) and values (monetary data)** contain the same amounts, or simply that quantity can be used as many were generated, or the production value obtained is equal to the sum of the result of the production and price of the amount of resources used.

The natural data and product/expenses (production, purchasing) price multiplication leads to the monetary data. Producer and purchaser prices are basically national currency and are presented in nominal terms, but for time series and spatial (between countries) comparability the SPEL BS includes the possibility of inflation adjustments, respectively in constant prices (for example, using the 1990 as base year) calculations.

Observing the current state of regulation and discipline of the agriculture and certain Euro area members, it seems clearly that without accurate accounts, without keeping all of the criteria, neither any system can be sustained, about which there are only partial information or “good estimates”.

In the case of columnar clearing, where one sector’s one average (nationally) units(e.g. 1 hectare or as one dairy cattle) are settled, that all the products and by-products accounted natural as well as the amount and the price element and their multiples (PV as Production Value), used inputs in natural and multiplied with (sometimes imaginary) price factors (e.g. ECU/kg or ECU90/kg) then the other expenses like the cost (fiscally VC, FC, TC) GM (PV-VC), than any aid, and finally the net Profit (NP).

3.2.1 Column oriented accounts

In case of SPEL a column/sectoral accounts for a vegetable sector:

- \+ main product(s), by-product(s), * price (ECU/kg)
- \- N, P, K fertilizer and manure (kg/ha) + lime * price (pl. ECU/kg)
- \- SEEP – seed (ECU90/ha), PLAP – plant protection (ECU90/ha) * price (pl. ECU/kg)
- \- REPV, REPO, ENEV, ENEO (ECU90/ha) * price (pl. ECU/ha)
- \- WATV, INPV, INPO – others * price (ECU90/ha)

- \+ PROV (ECU/ha)
- \- TOVA, TOOV, TOIN (ECU/ha)
- \+ GRMA, (PFSA, PFSB, PFSC), GVAM (ECU/ha)
- LEVL (1.000 ha)

Where: REPV: variable repair costs, REPO: fixed/overhead repair costs ENEV: variable energy costs, ENEO: fixed energy costs WATV: variable irrigation costs, INPV: other variable inputs, INPO: other fixed input factors, PROV production value (ECU/ha), TOVA: total variable costs (ECU/ha), TOOV: total fixed costs (ECU/ha), TOIN: total costs (ECU/ha), GRMA: Gross Margin (ECU/ha), PFSA-PFSC: Agricultural Policy subsidies (ECU/ha), GVAM: Net income, value added (ECU / ha), LEVL: national scale (1.000 ha).

The SPEL 12 letter codes are made up 3*4 letter members, where the first 4 letter term represents the sector, the second represents the factor in the sector (as main products, by-products and inputs) and the third member as BECB or BASB references the currency of the accounting, thus ECU (pre EURO clearing currency) in case of BECB and the national currency for BASB.

The categories/codes for controlling as PROV, TOOV, TOVA, TOIN, GRMA, GVAM, PFSA, PFSB, PFSC, MGVA codes are nothing more than the business economics classes (like PV as production value, GM like Gross margin, etc.) are listed in the 2nd paragraph above.

Here is 1 ha averaged soft wheat in Germany in 1992 as example (SWHE: Soft Wheat, namely wheat industry and its main product - ABTA Code) sectoral accounts:

2. Table Soft-wheat column oriented accounting in the SPEL (Germany, 1992, source: SPEL, own calculations)

SWHE			BECB				
SPEL code (1)	**Dimension. (2)**	**Value (3)**	**UVAL/Price ECU/ha**	**Multiplication (4)**	**SPEL code**	**SPEL value (5)**	**Check (6)**
SWHESWHEBECB	kg/ha	5990,79	151,531	907,791			
SWHESTRABECB	kg/ha	360,897	5,4	1,94884			
SWHESILABECB	kg/ha	0	11,723	0			
SWHEDHAYBECB	kg/ha	0	35,369	0			
SWHEPRADBECB	ECU/ha			0			
SWHECOWOBECB	ECU/ha	3,539	1000	3,539			
				913,279	SWHEPROVBECB	913,281	OK
SWHENITFBECB	kg/ha	131,663	493,927	65,0319			
SWHENITMBECB	kg/ha	75,466	246,964	18,6374			
SWHEPHOFBECB	kg/ha	30,546	541,316	16,535			
SWHEPHOMBECB	kg/ha	51,344	270,658	13,8967			
SWHEPOTFBECB	kg/ha	44,358	289,842	12,8568			
SWHEPOTMBECB	kg/ha	64,755	144,921	9,38436			
SWHECAOFBECB	kg/ha	163,666	28,984	4,7437			
SWHESEEPBECB	ECU90/ha	56,157	969,056	54,4193			
SWHEPLAPBECB	ECU90/ha	51,46	1100,251	56,6189			
SWHEPLOFBECB	ECU90/ha	16,206	943,922	15,2972			
SWHEREPVBECB	ECU90/ha	109,277	1104,859	120,736			
SWHEENEVBECB	ECU90/ha	37,502	1113,219	41,7479			
SWHEWATVBECB	ECU90/ha			0			
SWHEINPVBECB	ECU90/ha	54,62	1141,049	62,3241			
SWHEREPOBECB	ECU90/ha	8,661	1104,859	9,56918			
SWHEENEOBECB	ECU90/ha	35,857	1113,219	39,9167			
SWHEINPOBECB	ECU90/ha	3,585	1141,048	4,09066			
SWHEINADBECB	ECU/ha			0			
SWHEVATUBECB	ECU/ha			0			
				545,806	SWHETOINBECB	545,805	OK
				492,229	SWHETOVABECB	492,228	OK
				53,5765	SWHETOOVBECB	53,577	OK
				421,05	SWHEGRMABECB	421,052	OK
				367,473	SWHEGVAMBECB	367,476	OK
				0	SWHEPFSABECB	0	
					SWHEPFSBBECB		
				0	SWHEPFSCBECB	0	
				367,473	SWHEMGVABECB	367,476	OK

For livestock:

- \+ main product(s), by-product(s), * price (ECU/kg)
- \- FCER, FPRO, FENE, FMIL, FFSI, FDRY, FOTH, (kg/db) *price (ECU/kg)
- \- REPV, REPO, ENEV, ENEO (ECU90/ha) * price (ECU/ha)
- \- INPV, INPO (ECU90/ha)– others

- PROV
- TOVA, TOOV, TOIN,
- GRMA, (PFSA, PFSB), GVAM
- LEVL

Where: FCER: cereal fodder, FPRO: protein based fodder, FENE: fodder with high energy, FMIL: fodder from mill products, FFSI: silage, FDRY: dried fodder FOTH: other fodder, INPV: other variable input, INPO: other fix/overhead input, PROV: production value (ECU/ha), TOVA: total variable costs (ECU/ha), TOOV: total fix /overhead cost (ECU/ha), TOIN: total costs (ECU/ha), GRMA: gross margin (ECU/ha), PFSA-PFSB: agricultural policy subsidies (ECU/ha), GVAM: Net value added or profit (ECU/ha), LEVL: national size (piece, 1.000 db, etc.).

For livestock sector account as an example, a German dairy cattle (MILK - dairy cattle) column oriented accounting (products, expenses and profit) from 1996:

3. Table Dairy cattle/milk column oriented settlement/account from SPEL data (source: SPEL, own calculation/account)

MILK			BECB				
SPEL code (1)	Dimension (2)	Value (3)	UVAL/Price ECU/ha	Multiplication (4)	SPEL code	SPEL value (5)	Check (6)
MILKMILKBECB	kg/cow	5503,31	270,025	1486,03			
MILKBEEFBECB	kg/cow	80,622	2074,048	167,214			
MILKCALVBECB	calf/cow	0,863	90103,266	77,7591			
MILKDCOWBECB	cow/cow	0,735	717060,125	527,039			
MILKMANNBECB	kg/cow	85,75	246,687	21,1534			
MILKMANPBECB	kg/cow	44,739	272,73	12,2017			
MILKMANKBECB	kg/cow	104,391	132,207	13,8012			
MILKPRADBECB	ECU/cow		1000	0			
				2305,2	MILKPROVBECB	2305,157	OK
MILKFCERBECB	kg/cow	708,602	113,016	80,0834			
MILKFPROBECB	kg/cow	285,56	281,272	80,32			
MILKFENEBECB	kg/cow	699,11	93,951	65,6821			
MILKFMILBECB	kg/cow	286,596	89,626	25,6865			
MILKFDRYBECB	kg/cow	340,807	20,571	7,01074			
MILKFFSIBECB	kg/cow	16938,5	11,424	193,505			
MILKFOTHBECB	kg/cow	252,652	76,908	19,431			
MILKICOWBECB	db/cow	1	695926,375	695,926			
MILKIPHABECB	kg/ha	0	697,87	0			
MILKIPHABECB	ECU90/cow	105,167	1328,67	139,732			
MILKPLOFBECB	ECU90/cow	2,825	802,376	2,26671			
MILKREPVBECB	ECU90/cow	34,529	1269,069	43,8197			
MILKENEVBECB	ECU90/cow	25,96	1203,039	31,2309			
MILKINPVBECB	ECU90/cow	46,557	1328,67	61,8589			
MILKREPOBECB	ECU90/cow	6,054	1269,069	7,68294			
MILKENEOBECB	ECU90/cow	31,086	1203,039	37,3977			
MILKINPOBECB	ECU90/cow	2,446	1328,669	3,24992			
MILKINADBECB	ECU/cow			0			
MILKVATUBECB	ECU/cow			0			

				1494,88	MILKTOINBECB	1494,88	OK
				1446,55	MILKTOVABECB	1446,55	OK
				48,3305	MILKTOOVBECB	48,33	OK
				858,647	MILKGRMABECB	858,607	OK
				810,316	MILKGVAMBECB	810,277	OK
				0	MILKPFSABECB	0	
				0	MILKPFSBBECB	0	
				0	MILKPFSCBECB	0	
				810,316	MILKMGVABECB	810,277	OK

Whether we talk about animal or vegetable category, from business economics point of view, a proper accounting can be performed for all single units (hectare, head, 1000 pieces, livestock unit, etc.). It does not matter that we are talking about a farm, and it is made for board, or for sectors, or even county size. With the help of data warehouses, accounting could be expected theoretically even down to board/plot or the level of hectare everywhere.

3.2.2 Row oriented settlements/accounts

Contain changes in inventories for a single crop for the current year in farm, market and national level.

- yield (t/ha) * area (LEVL: 1.000 ha)
- Consumption within farm (PLOF, PCOM, PFEE, PSEE, PCSF),
- Market change: Import – Export ± PCSM – stock change in market
- Market „consumption”: PLOS, PCOM, PFEE, PSEE, PPRO)]

Where: PLOF: losses within farm, PCOM: consumption within farm, PFEE: animal feed within farm, PSEE: used as seed within farm, PCSF: stock change within farm, PCSM stock change in market, PLOS losses in market, PCOM: consumption in market, PFEE feed across market, PSEE seed across market PPRO: processing in the market. Unit: everywhere 1000 t.

The row oriented accounts as an example of the corn (maize) sector of Germany in 1997.

4. Table Corn's row oriented accounts from SPEL data for Germany 1997 (source: SPEL)

MAIZ			**BECB**		
SPEL code (2)	**Dimension (3)**	**Value (4)**	**SPEL code (2)**	**SPEL value (6)**	**Check (6)**
PINDMAIZBECB	1000 t	500,391			
PSEEMAIZBECB	1000 t	57,15			
PFEEMAIZBECB	1000 t	1160,804			
PCOMMAIZBECB	1000 t	1001,592			
PLOSMAIZBECB	1000 t	30,57			
		2750,507	PDOMMAIZBECB	2750,507	OK
PEXTMAIZBECB	1000 t	1077,5			
PCSMMAIZBECB	1000 t	-255,231			
PIMTMAIZBECB	1000 t	1853			
		1719,776	MAPRMAIZBECB	1719,776	OK
		1719,776	TRAPMAIZBECB	1719,776	OK
PCSFMAIZBECB	1000 t	-55,688			
FEEPMAIZBECB	1000 t	1457,81			
SEEPMAIZBECB	1000 t				
PCOFMAIZBECB	1000 t				
PLOFMAIZBECB	1000 t	66,103			
		3188,001	PROPMAIZBECB	3188,001	OK
MAIZMAIZBECB	kg/ha	8874,222			
MAIZLEVLBECB	1000 ha	359,243			
		3188,00213		3188,001	OK

3.2.3 Problematical fields

What is missing, or should be kept in any other way recorded?

- Changes in stocks like PCSM, PCSF are meaningless by itself, opening and closing stocks should exist too, in the absence of them, may not be able to know the opening and closing quantity at the beginning and at the end of the year of a given product. In principle it is calculable recursively, but when the opening data of the first year does not exit, than the recursive calculations neither work.
- The levels shall be distinguished between the producers, intermediaries, high users, and the state. For each layer the opening and closing inventory by years and mapping of all other movements should be required.
- PLOF, PLOS, category PPRO - (i.e.: processing and Processing Market, losses on Market, what are they)?
- SEEP: unit: ECU90/ha, the price of SWHESEEP (soft wheat seed) and MAIZSEEP (corn seed) is the same, as only UVALSEEP, PRICSEEP, PRIISEEP, and PRINSEEP exist because of the 8 letter long nomenclature. In reality two plant seed's prices are random the same for the same weight.
- The input positions commonly described with ECU90/ha units are describable by natural quantities too, and their consistency cannot be examined in this way, e.g. the seed: SEEP - PSEE and SEEP at formation/production side (row oriented) were in 1000t, at consumption side in ECU90/ha.
- NPK fertilizer and manner as forming product – shouldn't it recognized as product/stock?
- DEPR, ENEV, REPV – scaling the cost of machine use and scaling the depreciation of buildings for sectors.
- ENEO, REPO, INPO - overhead cost allocation, fixed-energy /fix repair expenses interpretation of them?
- The cost of tillage performed in contracted work, where to be shown: as INPV or after using calculations in ENEV, in REPV or in DEPR?

In the SNA-based EA so Σ (total areas$_i$ * Net profit$_i$, where i = crop sectors) would give the contribution of crop production for GDP/GNP. If under such a scheme (in the 1:1 mapping is impossible) we could account with all of the country's arable area (hectares per plot, per table etc.), it would be plausible, for example that the agriculture's share in GDP fell by around 3%. The concept was born in 1998 as MIMIR/ MIVIR / MITIR already and at that time all were then lined up behind: Ministry of Agriculture, ARH, AIK, CSO, ARI, Chambers of Commerce, Ministry of Finance, Tax Office, Customs, HMS, Product Advice, Chambers, Universities, HSA, OMFB, OTKA, PFP, FEFA, land registry, Mountain Villages, consultants, producers and professional organizations. However, it failed to take formed place later under any name.

Will be essential to establish a strict (comparing with the current) registration system - even over a limit, even in general terms. The SPEL scheme allows it with appropriate changes, for all participants to be able to settle with the produced goods and inputs. If this would give approximately the same results as adopted by the NAV tax declarations, it would be all right, but under the current rules of taxation, tax liability arises in case of sale of products or services are given in the background and they are billed or invoiced. In case of a not invoiced product movement – then no proof, nothing happened. The ratio of the not invoiced sales of good and given services etc. that is not "all the same", and above a certain ratio, it's no wonder that fell the share of the agriculture in GDP, which the post-transition period and more, the crisis since 2008 only strengthened.

3.3 Consistency and plausibility

The assignment/research task described in the methodology as solved task, context can be gleaned from the experience that the forecasted outlook is authentic when the evolution of output and their price - so to speak - are consistent (are in harmony) with the trend's "future values" fitted on the previous year's inputs and on their prices.

Under harmony/consistency is meant for firstly, that actually increase or decrease happened, beyond this the value the difference is not significant.

The resolved work's item number (number of plants) and the quantity of inputs used and their price were quite high (it was several days of work for four persons), so it's hardly would have said in case of one plant that it randomly happened, and on this base the conclusion is coated and described also.

Based on the earlier it can be stated:

- **First kind of consistency** (as well as guiding control criteria), the predicted phenomenon has a fully consistent and plausible value seen on the base of its input components or in its environment, namely the possible inputs and their environmental values are consistent with the expectable developments.

As a database like SPEL database containing so detailed definition of these production figures is rarely met, so that after the 2002 assignment, returning back to the available regularly updated databases as data resources, with less inputs supported degree of consistency was needed. It is none other than in connection with the ASM topics, some methods mentioned but regardless, the plausibility as a level/ benchmark of the validity.

The amount of a phenomenon can be assumed to be correct, if

- The level of the year/year change and the future value is not salient (within previous extreme changes – but for the extreme growth, sometimes there are actual examples.)
- The value (such as yield, price and as harvested area) is not negative.

On the base of the earlier:

- **Consistency of the second kind** (as well as guiding criteria for the controls - plausibility) when the phenomenon deemed or rejected appropriate in the light of its own value and its own past values.

Along the tests of forecasts occurred, that the widely used trend (more than once) gave a value (yield, price, area) which was negative, so in relation to the judgment of the correctness of the forecasted values, primarily because of comparability (!) was not an issue to exclude all wrong, and not plausible value.

The chronological weighted forecasts were excluded because of frequent negative forecasted values. This method's multivariate version called "Fundo_ chartist" counterpart has no problem with, because for forecasting one value, it uses 6 time-series, 5 data for each series, with the corresponding 5-5 weights (6*5 SUMPRODUCT) and 6 +1 parameters (the same parameter value for all six values and +1 parameter for each time series), in this solution does not really occur over amplitude and negative value. Enough fixed solution and so numerous independent variables (all of them should be close to the real value along the learning phase at the same time), occur almost as constraints.

Along the forecasts mostly (the similarity analyzes are based on longer periods) one value is forecasted on 5 years of data. Above all, the authenticity is which limits the long-term forecasts - although we had some study which showed, that for a longer period, e.g. for 3 years gave better results than for a short one-year term. The plant production would be enough to see the next one year – clearly and correctly – even than the eighth badly (see CAPRI).

3.4 Forecasts

3.4.1 On the base of direction hit

The goodness of the four forecasting methods by the primary criteria is included in the following Statement.

154		
155	attribútum	(mind)
156	adat_típus	találat_2
157	növény	(mind)
158	ország	(mind)
159	év	(mind)
160		
161	Átlag : érték	
162	módszer	Össz.
163	sine	64,70%
164	COCO	48,51%
165	fundo_chart_per_plant_per_year	49,03%
166	trend	41,74%
167	Végösszeg	50,99%
168		

1. Statement The average correctness of the methods for 158 time series on the base of direction hit (own collection)

For basic starting point can take the results of the linear trend, which gives 41.74% of bearing results out of the total 158 cases. For 50:50 ratio increase / decrease time series it causes 8.26% loss. In comparison, the wave approximation has been 64.7%, which is better (relative to the trend by 64%), but compared to the actual expected 75%, is even far. The other two, situated between 41.74% and 50% closer to 50%.

On the base of the ex-post test results it can be defined that, for a given crop / country combinations which was the best result. It was collected fully for yields, prices and area harvested in a statement. The best result of systematic selection is located under the following statement:

	Oats	16,67%	0,00%	33,33%	33,33%		33,33%
	Rye	16,67%	16,67%	16,67%	16,67%		16,67%
	Triticale	16,67%	33,33%	33,33%	16,67%		33,33%
	Wheat	16,67%	33,33%	50,00%	16,67%		50,00%
⊟ro	Barley	50,00%	50,00%	50,00%	33,33%		50,00%
	Maize	83,33%	50,00%	50,00%	83,33%		83,33%
	Oats	33,33%	33,33%	33,33%	33,33%		33,33%
	Rye	66,67%	33,33%	33,33%	33,33%		66,67%
	Sorghum	50,00%	16,67%	33,33%	50,00%		50,00%
	Wheat	66,67%	50,00%	33,33%	33,33%		66,67%
		64,92%	48,48%	48,92%	41,71%		
							73,14%
		wave	**COCO**	**fundo_chart_per_plant_per_ye**	**trend**		
						area harvested	**76,35%**
						price	**72,84%**
						yield	**70,26%**

2. Statement Results of the methods by attributes, by countries and by plants on the base of direction hit (own collection)

For the whole case collection 73.14% of correctness is obtained, which is close to the expected 75%. For the starting 41.71% point it is a significant improvement. Divided up to three, 72.84% (prices), 76.35% (area) and 70.26% (yield) come as partial result. Thus, in the case of the harvested area the 75% was also exceeded.

Rating the 73%, it means 3 right decisions out of 4, or in other words, one inappropriate decision can be a neutralized by a correct, than still have two, with "we produce" profit. Comparing with the random guessing's 50-50% (assuming that the actual ratio of change is 50:50, as decrease and increase) when profit is knocked out, there is no gain, than it's acceptable that we are still slightly ahead.

The result (s) evaluation in steps:

1. Firstly it should be pointed out, that the trend values (such a widespread method) was significantly exceeded in this way.
2. Ranking can be established between the methods, for the entire collection, or for parts too (price, yield, area). The rankings and the distribution for the total case-collection is the following (with tie-adjusted):
    a. In the ranking there's clearly the wave function approach on the first place (46.46%),
    b. on the second and third position there's the fundo_chartist (24.34%) and the similarity analysis (COCO) (19.47%),
    c. the trend is the fourth (9,73%).

sine	46,46%
COCO	19,47%
fundo_chart_per_plant_per_year	24,34%
trend	9,73%

In case of the methods for the wave-approximation alone, where some subjectivity could have been possibly present at selecting the parameters, at determining the test result. As specified in the methodology and in accordance with the first statement (64.92%), can be seen that not the "at any price to achieve the best results" was the aim so far, or from the log xls-s anyone can see that it would have plenty of parameter pairs - in many cases - which have significantly could have improved the wave function approach's results.

By determining the parameters the not very good learning (above 75-80% and usually below 94%) and the low summed difference were the two primary considerations. The result was selected by filtering on the base of the primary criteria and after it selecting a record (on the base of the second parameter), then the test value has shown itself.

For further evaluation of these resulted values the following realistically thinking shall be considered:

- It is difficult (if not impossible) to find such clear comparative values in increase/decrease (hit direction) question on any issue, in terms of competing institutions. Everyone considers itself a market leader and ranked on the basis of someone other's "independent*" measurements.
- Treating the hit direction ratio of the forecasts as efficiency matter, in terms of long averages we can be glad with 75% or with little more values. It may be desirable to get between 80% - 90% average - but maybe this won't go easily, and to get over 90% will remain only a dream.
- All methods (except Trend) can be seen as a learning system (such an Artificial Intelligence application - but that itself is a matter of intense orientation), which learn from past changes and calculate on the basis of the past anything in advance. A change that has never before occurred, there's no pattern for it, cannot be predicted. The wave method e.g.: example trying to catch/learn the continuous change of the trend so it is possible that after the downward may be a round up, then increasing, but there are samples when it happened in other way(s). Based on the calculated result it is still in front of the 3 other methods with 46% share.

3.4.2 Evaluation on the base of mean differences

Along analyzing the results, the following facts became clear. Over the direction hit, beyond the trend method, the others have greater differences between predicted and actual values.

The following table shows per forecast the ratio of the actual absolute difference and fair actual values in % form for 6 forecasts (attribute, country, and plant) stacked and averaged the difference %-es.

		Rye	23,26%	34,87%	15,95%	26,68%
		Triticale	83,58%	88,23%	61,40%	102,70%
		Wheat	86,18%	50,57%	56,22%	62,76%
	⊟ro	Barley	44,29%	24,27%	34,54%	23,33%
		Maize	47,09%	40,23%	45,50%	43,48%
		Oats	37,49%	27,48%	27,77%	31,48%
		Rye	20,40%	26,39%	20,04%	27,45%
		Sorghum	62,95%	38,41%	51,83%	41,73%
		Wheat	37,30%	32,17%	40,44%	51,61%
Sumtotal			20,39%	21,02%	16,61%	27,33%
			wave	COCO	trend	fundo_chart_per_plant_per_year
	mean of differences:		20,35%	21,01%	16,50%	27,09%
	st.deviation:		0,20	0,17	0,13	0,69
	st.deviation/mean:		96,11%	81,33%	80,10%	254,29%

3. Statement Average differences of the methods by attributes, by country and by plant for the 158 time series (based on own collection)

The trend has the lowest average deviation, while the Fundo_chartist method has the largest and is also true for the two extreme cases in case of the quotient of the standard deviation and the mean.

Ranked on the basis of differences, this sequence is given: viewed from lowest: trend – wave – COCO – fundo_chartist.

Since this was determined for all time-series, searching back for the winning method (taking into account the ties /

5. Table Tie adjusted average differences of the methods by attributes, by country and by plants (own collection)

	47,09%		45,50%	
	37,49%	27,48%	27,77%	31,48%
	20,40%			
	62,95%		51,83%	
	37,30%			
	20,02%	21,79%	16,39%	20,79%
	wave	COCO	trend	fundo_chart
mean difference:	20,10%	22,16%	16,39%	20,79%
st.dev.of differences	0,21	0,16	0,14	0,18
st.deviation/mean:	104,19%	74,27%	83,92%	87,37%

off-races) the average difference %-es will change (5^{th} Table):

Essentially the highest (fundo-chartist) average difference fell back, the COCO increased slightly, the other two are hardly changed, changed places in the rankings the former third and fourth place, but not with a significant difference.

Since the values of yield, price, harvested area scattered depending on the method, it is generally not worthy to generalize. The differences in % projected for 1 forecast in values are not so great, there are not as great differences based on the trend. The values are spread between 16.39% and 21.79%.

Since this is added aggregated to the 158 time series the dividing up for the harvested areas, for the price and for yields is the 6^{th} table:

Here the fluctuation for COCO, trend and Fundo_chartist methodologies can be seen, and the wave is fairly stable at the 20.02% in average. The item numbers orients in the sense that the ranking on the base of direction hit is clear and obvious in this order: the first position has the wave, on the second there is the fundo-chartist, then the COCO is the third and on the fourth place there is the trend. Since 68 cases of a tie, so the sample size was amended with 68, which is 158+68=226 cases.

6. Table Tie adjusted (for winner methods) average differences of the methods by attribute breakdown (own collection)

	wave	COCO	trend	fundo_chart_per_p	
mean difference (%) for 1 forecast:	**20,02%**	**21,79%**	**16,39%**	**20,79%**	
		harvested area			
mean difference (%)	20,52%	10,78%	5,56%	15,09%	
number of elements:	38	15	5	12	**70**
		price			
mean difference (%)	19,12%	27,76%	18,32%	25,09%	
number of elements:	37	21	6	11	**75**
		yield			
mean difference (%)	20,49%	26,78%	20,25%	21,46%	
number of elements:	30	8	11	32	**81**
	105	**44**	**22**	**55**	**226**
shares on the					

3.5 Attachment to the MTP and the Bayes' theorem

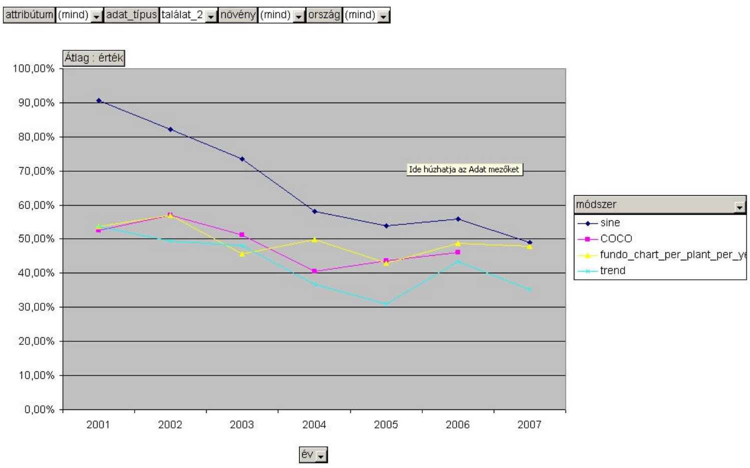
3.5.1 Attachment to MTP

In the MTP computations for the crop composition the gross margin is listed as coverage maker. The higher the value, the more it’s worth to take part from the planting plan - of course, up till that point if there’s another sector which utilizes a necessary resource better.

At the MTP-based approach the area is treated aggregated, although in reality it is started from the table (soil) and the forecrop to select the next crop, probably there is so many restrictive conditions inserted to be only those plants placed to which the forecrop is suitable. In relation to the ASM’s repeatedly been a strong tendency towards monocultures. Impossible to insert as many restrictions at regional level, and neither the resolution is possible for a single plot.

For the gross margin (GM) the question is that to determine the size of the current sector area (as endogenous variables) where the price and yield data from. The calculation is based on: [GM] = [Production Value] - [Variable Cost] and [Production Value] = [yield] * [Price]. VC (Variable Cost) is assumed to be known.

The yield and price predictions (right method for the right plant and country), according to the measurements work with 70.26% and 72.84% direction hit ratio. The 4 method could be adapted to all data of each table/plot, after identifying probabilities preparing plan for crops.



5. Diagram Direction hit ratios by years and by methods (own collection and representation)

Unfortunately, the Mészáros-Csáki book doesn't writes about that as a result of the (dynamic) linear programming method, the modified sowing plans, caused how much more profit (GM, Net Income, Profit). Unfortunately, the book is read with an impression that the state-farms accepted the proposed sowing plan and in the first year it was implemented. Then there is no news about how better became anything with or without optimization. Examples are included, but no concrete results in this topic.

The MTP and other nonlinear optimization methods/solutions (see ASM chapter and e.g.: GE models, the tendency towards monoculture) in normal case the total area is seen aggregated and the method distributes the area for different plants on the base of income-generating capability. The constraints on farm size is none other than the possible limiting factors (e.g.: machine capacity) and can be specified indirectly by the crop rotation restrictions. In national scale there is no way to include so many conditions there is always some tendency toward monoculture. The reason is not else that we approach not from the production unit, or from a table or plot level, but from an aggregate point of view.

3.5.2 Attachment to the Bayes' theorem

The Bayes' theorem shows the ex-post manner, the joint occurrence of two events on defined frequency basis, shows the probabilities of the method of use. Apart from the occurrence of two events, the 4 method's forecasted values suitability verification is not else than the ex-post way defined predicted probabilities. Thus, the Bayes theorem (decision theory) relevant parts section:

> *If we take into account the fact that the z1, z2 and z3 states z1 = 0.2400, z2 = z3 = 0.4600 and 0.3000 probabilities occur, the maximum expected values of the lines may be weighted in Block 4 with these probabilities, and results from the sum of the long-term average gross margins can be calculated, with the consequent use of the April weather information.*

From this, mostly that "may be weighted with probabilities" is, which is the same, as the probabilities (of occurrence) of the predicted values. This is so because, that e.g.: the wave method out of the 158 cases, by 105 times is the winner, and in this cases the average right direction hit is 0.74. The mean difference was 0.20 (for 105 cases). When the forecast predicts an increase, then with an average (!) of 20% gives a value above the actual value. This is 1+0.2 = 1.2, weighting this 1.2 with 0.74, 0.89 is obtained. The closer this value is to 1, the more accurate each/all predicted value in average. The real problem here is the non-uniform distribution of the time-series and forecasted values, but this is present at each method, and this problem seen as "aggregated" in case of 158 time-series, cannot be resolved.

In case of decrease it can be said that weighting is correct, just in this case, shall be weighted with reciprocal value. In case 20% of underestimation that is 1-0.2 = 0.8. That's worth not to multiply, than divide it by 0.74 to gain a value of around 1.

The relevant data for the four methods can be found in the following table:

7. Table The average accuracy of the forecasts after weighting (own calculation)

	wave	**COCO**	**trend**	**fundo_chart_pe**
mean difference:	0,20	0,22	0,16	0,21
average hit direction of forecasts:	0,74	0,71	0,69	0,73
case of increasing forecasts:	**0,89**	**0,86**	**0,80**	**0,88**
average accuracy				
(1 is the accurate)				
case of decreasing forecasts:	**1,08**	**1,11**	**1,21**	**1,09**

The order comes in the following way: wave fundo-chartist, similarity analysis (COCO) and trend. Fortunately, the order is completely mirrored in the case of ascending and descending projections, so the methods are (accidentally) pretty consistent.

3.6 Unified statistical data collection

Referring back to those described in connection with SPEL:

> „*Will be essential to establish a strict (comparing with the current) registration system - even over the limit, even in general terms. The SPEL scheme allows it with appropriate changes, for all participants to be able to settle with the produced goods and inputs.*"

In general, most of all the unified and above a certain size the extended data collection is, which necessity shall be emphasized. Uniformed data collection in order to respect all participants (politicians, chambers of commerce, farmers) to report the same phenomenon in the same units. With this "standardization" all phenomenon will be measurable and comparable at country-, region-, county level.

3.7 Additive parallelism

Factors are additive in the statistical time series analyses (ARMA and ARIMA models), in the chronological weighting forecasts, in the fundo-chartist forecasts, at the DEA multiple (and at it's all other) form, and at similarity (COCO) - based forecasts. In general it`s known that input factors are not exclusively used additively, there are some which limits the rest/other. So to achieve the ideal production function (e.g. DEA, DEA simulation multiplier formula) with a purely additive interconnection of functions is excluded.

In contrast, only the additive method that is in use. Attempts have been made is made and will be made with only multiplicative, and mixed (additive and multiplicative) methods, but not with breakthrough success.

4 CONCLUSIONS AND SUGGESTIONS

The primary issue raised in the introduction, over the tightening the document discipline, making the SAPS area/table based accounting/settlements required, which may arise by SPEL scheme or on a modified SPEL scheme.

Forecasts - if their direction hit and accuracy are acceptable, could be used with ARI(AKI)'s spring and autumn report's background information, it would be eligible for next year expectable harvested area, yield and price per plant which could serve for the base of any intended decision. With these planned decisions, become possible to achieve the quantity and price stability, at least within the borders.

Below is a list beginning with the operational level (1st point), and over it from the 2nd point down the strategically achievable levels:

1. Requesting yields for each physical block/plot: all SAPS claims should contain the quantities of crops harvested posteriorly. Then would be calculable on the basis of existing yields and PIR / PAIR (price) data and on the basis of the estimated sector (FVM's Technological Institute) costs:
    a. plot area based value added (GDP) values,
    b. net income and value added (GDP) ratio could be controlled from the basic GDP values
2. Arises the possibility of personalized predictions based on the requested area yields or plant proposal on the base of forecrop, on the surface of the SAPS claims, consultant function integration on Scandinavian sample.
3. area optimization, planting plan suggestion (per farmer),
4. sector efficiency calculations per table (compared to national averages based on the given input data) with the simulated DEA method,
5. Over time, improved forecasts opportunities for all areas.

In 2013, it is regrettable to see that from 87 thousand registered producers only those who are in the representative sample shall provide any production data for the CSO and for the Agricultural Research Institute, from which "pretty good estimates" can be made, but they still remain estimates, rather than descriptive facts for the entire agriculture. It can be said, that the Common Agricultural Policy (CAP) provides the framework, but if any of these frames, do not ensure the stability of the existing system, causing continued instability, throws, combing, then perhaps it would be the best to prepare for these.

The SPEL schema could be the primary step towards creating standards, therefore, the most important proposal over the yield data at plot level as described in point 1, to base and create the agricultural data asset management.

1. The Hungarian government shall have worked out the OSAP (National Statistical Data collection Program) equivalent data asset-management principles for the Hungarian Agriculture with the experts of CSO, ARI, and connecting research institutes, in particular, the operational and strategic quality assurance methodologies (see SPEL, elaboration of the principles of adaptation of SPEL) and reporting obligations, report- interpretation (decision-preparation) standards, etc. are necessary that our country would be able to recognize the first signs of unsustainability in a preventive, proactive and sovereign way on the basis of the statistical data collections ordered by the law and take the necessary steps to protect against these unwanted phenomena.
    a. Interval: 12 months,
    b. Responsible for: ad hoc appointed data asset manager, commissioner,
    c. Resource requirements: 1 billion HUF, (based on MIMIR, IACS, RIIR and other prior project's estimated and actual costs)
2. On the basis of the finished development described in paragraph 1 becomes possible, to deal with listed activities in point 2 such as prediction, design, optimization steps in any system, so this is possible only after the implementation of real data asset management.

5 SUMMARY

The first part of the thesis is the replacement of a widely known efficiency calculation method (DEA) which is a slow, time-consuming process. The replacement was the goal and the result. To be effective, as the number one economic interest is present at all times, so it is not possible to go around. The own work was the reuse of the original equations from DEA, which is true, that there are more variables to work with, but it does not need to be n * 10.000 or n* 100.000 times (n is the number of production units) recalculated, so in daily life, it can give a guaranteed faster solution.

In the next section the SPEL database and the SPEL logic based settlements should be considered as creating standards (obviously with their adaptation), as everything could be measurable and collected data by this system (down till board/plot level) could be the basis for the next section, consistent forecasts and the rational planning based on the forecasts. The SPEL accounting logic itself is clearly the result of EUROSTAT and ASA/IAP Bonn cooperation, but the contained data asset to extract into row and column direction settlements/accounts no one watched or checked (universal for all countries, for each year, for each sector combination) whether they correct are. There was a significant development covering the EU and the accessing countries in 2004, with resulting databases, accounts and simulation models, nobody reviewed/validated them until the summer of 2000 and 2001. This happened in the summer of 2000, and from the current run of the settlements/accounts it is clear that at home (in Hungary) the resulting databases are not arbitrarily accurate or closed. It is therefore necessary to establish the standards under either SPEL, or under as amended/modified SPEL logic, that the accounts shall be closed for each production activity in natural and in monetary in all respects. From this work the result of the own work is, that after expanding the data and programming down the accounts from SPEL, it is clearly argued and stated, that after the general introduction of this accounting logic, only dramatic transparency increase that all stakeholders may face to, due to the problems mentioned in the introduction, is essential would be. With the accounting logic of self-enclosed and self-checked inside, with this knowledge an adapted SPEL would be creatable also for ease introduction.

The most weighted literature part, the agricultural sector models wanted to show that these systems describing all sectors can be used well, independently from their internal model approaching realities (static, dynamic, equilibrium, dis-equilibrium etc.), if the results of these models as the modeled /simulated (or scenario run) returned the real “future”. As this is absolutely no evidence (which in turn we saw the opposite e.g.: in CAPRI simulation run for 2013), so the obvious solution is to be able to give/forecast the exogenous future values. So basically an agricultural sector model with a good internal logic, with good exogenous input data may well model the expected events, but it should be validated before ex-post manner. The zero step, before the model validation, is the exogenous values prior ex-post validation. This work was presented in the second part of so called "consistent and plausible" forecasting. It is desired to point out that in crop production the so-called stochastic processes are not necessarily unrecognizable, regarding the expectations of decision-making processes, in accuracy of information requirements, as long as approximately 75 % direction hit (as a new result) value is acceptable.

This part of the work, which should built into every design process among public and market participants and should use on the base of it by occasionally consultations prior to anticipated events. In the 90's, the Agricultural Market Organization (ARH) and it’s institutions (e.g.: AIK) was created for the same purposes, and then from 2004 to the former institutions integrated into the ARDA (MVH) which since the 1st May 2004 onwards, - since the beginning of the EU "doctrine" - performs implementing, monitoring, and more specially the paying agency functions.

Before the EU accession period the typical "problem-solving" method was the so called "bomb disarming", but it was not a lonely case, and for all of this type of "problem solving", the subsequent crisis management is the best description. To prevent it, these forecasts would be appropriate to prepare for. According to the tests, sometimes 75% or with even better direction hit, but in order dialogue, negotiation and cooperation is needed between the interested parties, both at national and EU level, at appropriate intervals, or in urgent cases on an ad hoc basis, but not subsequently.

The negotiations shall be based on the data of the agricultural production (in case of crop production) describing system, down to plot level, that is, a data warehouse/database may be the primary goal, which would give the possibility to handle its data-fortune, and the cause and aim oriented use of this data fortune.

Other improvements may be with assigning the data to plot level, thus area-based forecasts, plant suggestion on the basis on the forecrop, than area optimization for individual farmers and after these optimization runs national sowing plan could be made, which if is available in time, would leave time for all stakeholders - mainly for the State/EU - to prepare.

6 NEW AND RECENT SCIENTIFIC RESULTS

In the presented work, the following concrete results can be named.

1. Managed to resolve the DEA efficiency calculation procedure, which operates with the original so called multiplier form solution, where the sum of the efficiency values should be maximized in one solution run (instead the original solution, with numerous goal-functions), which is more stringent compared to the original, but much simpler and faster to run.
2. Expanding the SPEL database, programming column and row oriented accounting universally. The SPEL as framework (data structure) provides the possibility to plot/table level accounting self-mechanisms. The MEPÁR browser has been made for SPS/SAPS application registration support, a graphical (GIS based) system. Completing it with standardized row and column oriented accountings at plot level, would give the chance for following the agricultural activities exactly in physical production (quantity) and monetary (added value) value within the country borders.
3. Consistency:
   - Primary (full) consistency: the forecasts drawn from the inputs and their prices (in 2002, trend-based) highly correlated with the real values (were in the same direction and they were in tight scatter) of the outcomes produced and their prices. –Thus, there are no coincidences, the more input will likely cause more outputs as well, if the inputs are more expensive, then the outputs will be more expensive too.
   - Secondary (partial) consistency: the given phenomenon can be measured relatively to itself and its value. It is thus a form of plausible state, when the value can be graded on the base of the predecessors. It is within a meaningful value (interpretation range), and the assumed change is acceptable also.
4. Forecasts: The 4 methods (trend, similarity analysis, wave, fundo-chartist) were systematized; the tests were carried out on the same database for 1year predictions. The linear trend, as commonly used method (year-on-year basis, growth, decline or stagnation hit) could took only 9.73% share, out of the 100%, as the initial assessment method. The wave-function approximation proved to be the best out of the four with 46.46% share result. During the tests the systematization itself is a significant achievement, as with its help can be given by country and by plant which method’s use can be imposed or recommended.
5. With the correct predictions may become any kind of design method to a value added providing device, as without the corner points (price, yield, area), there is nothing actually to optimize to. Weighting the forecasted values with the ratio of the average direction hit results, closes the under or above estimated values to the actual value.

7 PUBLICATION LIST

7.1 Articles in scientific journal in foreign language

1. Bunkóczi László, Dr.Pitlik László, The state of agricultural sector modelling and possibilities of further development based on similarity, Gödöllő (SZIE Bulletin), SZIE, Gödöllő, 2007, 103-116, 14 p., ISSN 1586-4502
2. Bunkóczi László, Agricultural sector models and forecasting, NYME, Sopron, Gazdaság és Társadalom, 2013, 14 p., under publishing

7.2 Articles in scientific journal in Hungarian language

1. Pitlik László, Bunkóczi László, Pásztor Márta, Pető István, Popovics Attila, Online és lokális döntéstámogatási modellek fejlesztési lehetőségei és várható hatásaik, Szigma, Gazdaságmodellezési Társaság, Pécs, 2002, 169-175, 7 p., ISSN 0039-8128
2. Bunkóczi László, SPEL alapú ágazati elszámoló rendszer elvi sémája valamint konzisztencia ellenőrzött előrejelzések, ACTA AGRARIA Kaposváriensis, Kaposvár, KE-ÁTK, Kaposvár, 2011, 13 p., ISSN: 1418-1789

7.3 Presentations held in foreign language and published in scientific conference proceedings

1. Pitlik László, Bunkóczi László, Pető István, Creating objective evaluations, or the possibilities of similarity analysis, 19th WFSF Conference, Corvinus Egyetem, Bp., Corvinus Egyetem-WFSF, Budapest, 2005, , 5 p.,
2. Pitlik László, Bunkóczi László, Pásztor Márta, Pető István, Popovics Attila, Szűcs Imre, Consistency controlled future generating models, EFITA & WCCA'2005, Vila Real, Portugália, EFITA-APDTICA-UTAD, Vila Real, 2005, 1037-1044, 8 p., ISBN 972-669-646-1
3. Pitlik László, Bunkóczi László, Pásztor Márta, Pető István, Popovics Attila, Pisartsov Andrej, eFARMER - CAP-related and Country-specific eContent Management, EFITA & WCCA'2005, Vila Real, Portugália, EFITA-APDTICA-UTAD, Vila Real, 2005, 470-477, 8 p., ISBN 972-669-646-1
4. Pitlik László, Bunkóczi László, Pető István, Poster: Case study: Benchmarking an EU-member and candidate countries on the base of COCO, IAMO Halle Workshop, 2004, IAMO, Halle, 2004, , 2 p.,
5. Pitlik László, Pásztor Márta, Pető István, Bunkóczi László, Wichtige Aspekte der Dientsleistungsstandards für Anbieter landwirtschaftsrelevanter online Wissenskomponenten im Bereich e-government, GIL-Jahrestagung 2003, GIL, Berlin, 2003, 108-111, 4 p., ISBN 3-932987-05-5
6. Pitlik László, Pásztor Márta, Pető István, Bunkóczi László, Integrated Administration and Control System in Ungarn: auf dem Weg zu einem flexiblen und effektiven System, 23. GIL-Jahrestagung,Dresden, 18-20. September 2002., GIL, Drezda, 2002, 162-165, 4 p., ISBN 932987-04-7
7. Pitlik László, Bunkóczi László, Pásztor Márta, Pető István, IACS und Betrieb in Ungarn, IAMO Halle Workshop, 30. September 2002, IAMO, Halle, 2002, , 1 p.,
8. Pitlik László, Bunkóczi László, STOCKNET, avagy Hasonlósági függvények elemzési célokra, EFITA, Bonn, EFITA, Bonn, 1998, , 4 p., ISBN 3 932887 07 7

7.4 Presentations held in Hungarian language, and published in scientific conference proceedings

1. Bunkóczi László, SPEL alapú ágazati elszámoló rendszer elvi sémája valamint konzisztencia ellenőrzött előrejelzések, Kaposvár, VIII. Alkalmazott Informatika Konferencia 2010, KE-ÁTK, Kaposvár, 2011, 13 p., ISSN: 1418-1789
2. Bunkóczi László, Dr.Pitlik László, Szűcs Imre, Pető István, Konzisztencia-vezérelt modellezés lehetőségei, avagy szögfüggvények alkalmazása vetésterület, hozamok és árak előrejelzésére

mint GPS eljárás, Gödöllő (Tradíció és Innováció Konferencia), SZIE GTK, Gödöllő, 2007, , 9 p., ISBN 978-963-9483-85-9

3. Pitlik László, Bunkóczi László, Pető István, Környezeti-ökológiai konzisztenciák bevonása a modellalkotás automatizálásába, Biometriai Konferencia, Corvinus (KEÉ), Corvinus-KÉE, Budapest, 2005, , 7 p., ISBN 963 218 733 4
4. Pitlik László, Bunkóczi László, Pásztor Márta, Pető István, Popovics Attila, IDARA & IACS, avagy úton a modern adatvagyon-gazdálkodás felé, Agrárinformatika'2002, Debrecen, DE-ATC-MAGISZ, Debrecen, 2002, , 8 p., ISBN 963 204 869 5
5. Pitlik László, Bunkóczi László, Pető István, Agrárszektor-modellek az IDARA projekt tapasztalatai alapján, Gazdaságmodellezési Konferencia, Balatonfüred, Gazdaságmodellezési Társaság, Balatonfüred, 2002, , 14 p.,
6. Pitlik László, Bunkóczi László, A DEA módszer felhasználási lehetőségei az üzemhatékonyság mérésében, Agrárinformatika'99, Debrecen, DATE-MAGISZ, Debrecen, 1999, 313-328, 16 p., ISBN 963 7177 94 9

7.5 Research reports

1. Pitlik László, Bunkóczi László, Comparative analysis of agricultural policies by FAPRI, OECD and IDARA forecasts in the case of Hungary for 2006, IDARA jelentés, MIAU, Gödöllő, 2003, , 10 p., ISSN 1419-1652
2. Pitlik László, Pető István, Bunkóczi László, Trend-alapú szakértői vélemények generálásának automatizálása agrár-szektormodellek exogén változóinak prognosztizálására, IDARA projekt részeredmény, MIAU, Gödöllő, 2003, , 6 p., ISSN 1419-1652
3. Pitlik László, Bunkóczi László, Pásztor Márta, Pető István, Popovics Attila, OTKA zárójelentés, OTKA F030664 zárójelentés, MIAU, Gödöllő, 2003, , 16 p., ISSN 1419-1652
4. Pitlik László, Bunkóczi László, Pásztor Márta, Pető István, Popovics Attila, MI-alapú prognosztikai módszerek az EU és Mo. agrárstat-i adatbázisain alapuló előrejelzésekhez, OTKA 2.részjelentés, MIAU, Gödöllő, 2001, , 20 p., ISSN 1419-1652

7.6 Other

1. Pitlik László, Ónodi Bertalan, Pásztor Márta, Popovics Attila, Bunkóczi László, Wojcicka Ivett, Andrzej Wojcicki, Tanulmány a MIMIR megvalósulásáról, Magyar Integrált Mezőgazdasági Információs Rendszer - Tanulmány, MIAU, Gödöllő, 2003, , 35 p., ISSN 1419-1652
2. Pitlik László, Bunkóczi László, Pásztor Márta, Pető István, Popovics Attila, E-kormányzat projekt zárójelentés, MEH - E-Kormányzat Tanulmány, MIAU, Gödöllő, 2003, , 141 p., ISSN 1419-1652
3. Pitlik László, Bunkóczi László, Pásztor Márta, Pető István, Popovics Attila, E-kormányzat projekt 12.rész, MEH - E-Kormányzat Tanulmány, MIAU, Gödöllő, 2003, , 8 p., ISSN 1419-1652
4. Pitlik László, Bunkóczi László, Pásztor Márta, Pető István, Popovics Attila, Konzisztencia-vezérelt jövőmodell, avagy az agrár-szektormodellek új generációja, OTKA 2004-2007 pályázati anyag, MIAU, Gödöllő, 2003, , 38 p., ISSN 1419-1652
5. Pitlik László, Bunkóczi László, Pető István, FROM DATA TO DATA MINING…, Giessen-Gödöllő Szimpózium'2002, MIAU, Gödöllő, 2002, , 19 p., ISSN 1419-1652
6. Bunkóczi László, Hagyományos és mesterséges intelligencia alapú tőzsdei prognózisok és Autonóm Adaptív Ágensek, ETDK, Gödöllő, GATE GTK, Gödöllő, 1998, , 80 p.,