

Implementing endogenous technological change in a global land-use model

Christoph Schmitz · Jan Philipp Dietrich ·
Hermann Lotze-Campen · Christoph Müller ·
Alexander Popp

Presented at the 13th Annual Conference on Global Economic Analysis, Penang, Malaysia, 09/06/2010

Abstract Technological change in agriculture plays a decisive role for meeting future demands for agricultural goods. Especially in the longer run, i.e. several decades, technological change will be one of the major determinants of agricultural production. However, up to now, most agricultural sector models and models on land use change have used technological change as an exogenous input due to various information and data deficiencies. This paper provides a first attempt towards an endogenous implementation based on a measure of agricultural land-use intensity, called τ -factor. We relate this measure to empirical data on investments in technological change as well as production costs. Our estimated yield elasticity with respect to research investments is 0.24 and production costs per area increase linearly with an increasing yield level. Having implemented this approach in the global land-use model MAGPIE ("Model of Agricultural Production and its Impact on the Environment") we are able to project about required yield growth rates in the future. Highest future yield increases are obtained for Sub-Saharan Africa, the Middle East and South Asia. A validation with FAO data for the period 1995-2005 shows that our model behaviour is in line with recent observations.

Keywords technological change, land-use, agricultural productivity, land-use intensity, infrastructure investments, research and development

Christoph Schmitz and Jan Philipp Dietrich contributed equally

Christoph Schmitz · Jan Philipp Dietrich · Hermann Lotze-Campen · Christoph Müller · Alexander Popp
Potsdam Institute of Climate Impact Research (PIK), Telegraphenberg A 31, 14473 Potsdam, Germany.

Christoph Schmitz
Department of Agricultural Economics and Social Sciences, Humboldt University, Philippstr. 13, 10115
Berlin, Germany. E-mail: schmitz@pik-potsdam.de

Jan Philipp Dietrich
Department of Physics, Humboldt University, Newtonstr. 15, 12489 Berlin, Germany. E-mail: dietrich@pik-potsdam.de

1 Introduction

More than 200 years ago Thomas Malthus published his rather pessimistic population essay, which stating that population growth would be restricted by a low growth rate of food production (Malthus, 1798). Now the world is inhabited by almost seven billion people, which marks an increase by about 600% since Malthus' times. One of the main shortcomings of his essay was the underestimation of technological change (TC - as defined in Table 1) in agriculture (Trewavas, 2002).

concept	description
agricultural land-use intensity	degree of yield amplification caused by human activities (Dietrich et al., 2010)
τ -factor	measure proportional to agricultural land-use intensity (Dietrich et al., 2010)
technological change (TC)	more efficient usage of the input factors land, labour or capital (Romer, 1990)
TC investments	composite of annual investments in R&D and infrastructure (e.g. transport and telecommunication) [US\$/year]
investment-yield ratio (IY ratio)	TC investments required per human-induced unit yield growth and area [US\$/ha]

Table 1 Concepts and terms used in this paper

However, during Malthus' times technological change was negligible and higher food production was almost exclusively due to an increase in production factors (Federico, 2005). Important innovations in agriculture from the 19th century onwards changed this pathway (Runge et al., 2003). Since then land-saving technological change has been the main driver for growth in agricultural output (Wik et al., 2008; van Meijl and van Tongeren, 1999; Rosenzweig et al., 1988). Figure 1 shows the strong correlation between agricultural output and population during the last 200 years. Agricultural output has increased considerably, paving the way for strong population growth. Most of such increases in agricultural output have been the result of technological change induced by investments in Research & Development (R&D). One example is the so called "Green Revolution" in Asia and Latin America, initiated by two international agricultural research institutes (Evenson and Gollin, 2003)¹.

While the importance of TC in agriculture is widely acknowledged in the recent literature (Alston et al., 2009; Huffman and Evenson, 2006; Alene and Coulibaly, 2009; Thirtle et al., 2003), in most agricultural sector models and models of land-use change, TC is still implemented as an exogenous driver (Schneider and Schwab, 2005; Heistermann et al., 2006; Verburg et al., 2009). In these models, projections primarily depend on a fixed technology path rather than on internal model dynamics. This may lead to serious biases in model results due to an underestimation of the adaptability in the agricultural sector, especially in the longer run. In this paper we present a first attempt of implementing endogenous technological change in a land-use model, which means that the model can freely decide about the amount of technological change in the future.

The main reason for using an exogenous TC path in most models is that although the relationship between R&D investments in agriculture and technological change is well doc-

¹ During the 1960s and 70s the International Maize and Wheat Improvement Center (CIMMYT) and the International Rice Research Institute (IRRI) developed high-yielding wheat and rice seeds.

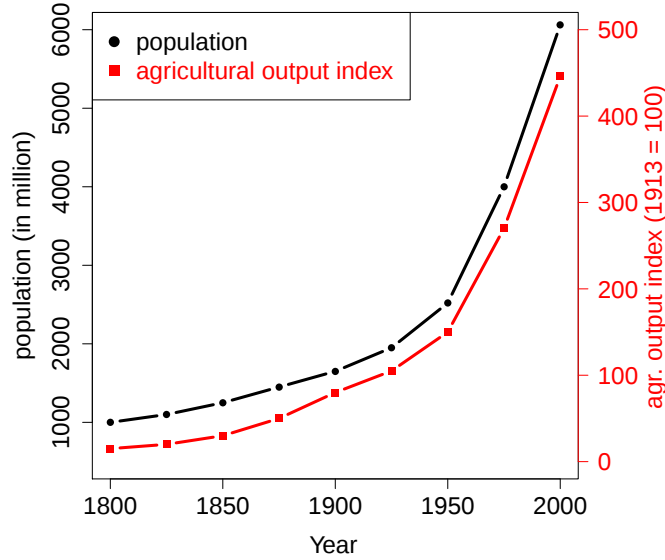


Fig. 1 Historic development of agricultural production and population [own illustration based on Federico (2005) and UN (2005)]

umented (Alston et al., 2009; Huffman and Evenson, 2006; Alene and Coulibaly, 2009; Thirtle et al., 2003; Federico, 2005; Pardey and Craig, 1989; Alston, 2000), the exact influence of R&D on technological change is still unknown. Several reasons exist for this knowledge gap. First, available time series of R&D investments are still relatively short (less than 30 years) and often incomplete (Pardey and Beintema, 2001). Second, spillover effects hamper the correct assignment of R&D investments to their impact. Third, success in R&D is hard to predict. High investment may fade away without producing any output, whereas in other instances low investment may create marvelous results. Finally no clear boundary exists between R&D investments in different sectors. In many cases inventions in one sector are based on inventions in other sectors. In a sector analysis of a specific R&D sector, e.g. agricultural R&D, these cross-connections cannot be considered.

In order to deal with these information deficiencies, we have developed a new approach which relates investments in technological change and corresponding yield growth to the τ -factor, a measure of agricultural land-use intensity (Dietrich et al., 2010)). The use of cross-sectional country data in combination with the τ -factor enables us to take the land-use intensity of countries as a surrogate for missing time series data. The problems of high uncertainty and unpredictable rates of return associated with investments and the problem of spillovers are partially compensated by using a high aggregation level of only ten world regions².

² AFR = Sub-Sahara Africa, CPA = Centrally Planned Asia (incl. China), EUR = Europe (incl. Turkey), FSU = Former Soviet Union, LAM = Latin America, MEA = Middle East and North Africa, NAM = North

In addition, we can also empirically show that the level of agricultural production costs per area evolves with the τ -factor. Implementing both aspects in a land-use model allows for computing an endogenous pathway of technological change, specifying required TC investments and changes in production costs. In this paper we use the land-use model MAgPIE ("Model of Agricultural Production and its Impact on the Environment") (Lotze-Campen et al., 2008, 2009; Popp et al., 2010) for the implementation and validation of our method. We provide a future agricultural technology pathway that can be used in other land-use change projections that require external assumptions on technological change.

2 Methodological framework

The internal computation of agricultural technological change is based on the production costs and the effectiveness of R&D investments on yields (investment-yield ratio, IY). The IY ratio evolves with the agricultural land-use intensity, describing the TC investments required per unit of yield growth and area. Accordingly, the production costs are based on the agricultural land-use intensity.

2.1 Measure of Agricultural Land-Use Intensity

We start by looking for an appropriate measure for agricultural land-use intensity. Dietrich et al. (2010) defined an index for agricultural land-use intensity (the τ -factor). It is derived using yield data from FAOSTAT (FAOSTAT, 2009) and yield data from the global dynamic vegetation and crop model LPJmL (Bondeau et al., 2007). Since the measure is output-related, it captures the full spectrum of yield increasing technology and management options.

$$\tau(c, i) = \frac{Y_{act}(c, i)}{Y_{ref}(c, i)} \quad (1)$$

The τ -factor is the ratio of actual or observed yield Y_{act} divided by a reference yield Y_{ref} which would be observed under a well-defined level of human activity (equation 1). It measures the amplification of yield due to human activities. In the case of yield growth which is not caused by natural changes in the physical environment (such as yield reductions due to water shortage in a dry year), the growth in τ is proportional to the growth in yield. Hence, it can also be interpreted as a yield index which is independent of natural physical conditions. The τ -factor is calculated for every crop type c and world region i .

2.2 Investment-Yield Ratio

Based on this measure of agricultural land-use intensity, we can now relate costs to technological change. For this purpose we have identified two types of costs which mainly influence the rate of technological change: first, public and private investments in agricultural R&D, and second, investments in infrastructure (e.g. transport and telecommunication). Data for public and private R&D investments are taken from IFPRI for the year 1981

America, PAO = Pacific OECD (Australia, Japan and New Zealand), PAS = Pacific Asia, SAS = South Asia (incl. India)

(Pardey et al., 2006) and data for infrastructure investments are from the GTAP database, version 7 (Narayanan and Walmsley, 2008) (discounted from 2004 to 1995)³. We divided both by the average yield growth rate observed in the years 1990-1999 taken from FAO (FAOSTAT, 2009). The reason for taking the R&D investment data of the year 1981 is the typical time lag between investment in R&D and its impact. The literature offers quite a wide range of various delays and lag-structures proposed for agriculture, ranging from a few years to several decades (Pardey and Craig, 1989; Alston et al., 1998; Fan et al., 2002; Cox et al., 1997). We chose a delay of 15 years, which is approximately the average of the delays used in the literature and, according to Alston et al. (1995) and Alston (2000), the time which is needed to reach the maximum value of gross annual benefits.

As a result, we get for each region the relationship between investments in agricultural research and the associated yield growth 15 years later. However, the absolute size of investments still depends on the size of a region: the bigger the region, the higher the variation in physical conditions. As a consequence, more research is needed to produce the same average growth rate compared to a smaller region with less variation in physical crop conditions. Consequently, we normalized investments relative to the agricultural area of a region. Specific R&D investments per unit of yield growth are computed as the ratio of R&D expenditures per area and the yield growth 15 years later. We apply the same concept for infrastructure investments, except that we assume no time delay. Both components add up to the investment-yield ratio IY describing the TC investments per area required per human-induced unit yield growth.

To relate this IY ratio to the τ -factor we have calculated the elasticity ϵ_{τ}^{IY} , i.e. the proportional relationship between an increase in the τ -factor and an increase in the IY ratio.

$$\frac{dIY(\tau)}{IY(\tau)} = \epsilon_{\tau}^{IY} \cdot \frac{d\tau}{\tau} \quad (2)$$

The elasticity ϵ_{τ}^{IY} is estimated via a regression analysis. Since agricultural R&D data are generally aggregated over all agricultural sectors and spillovers are expected, we have used an aggregate version of the τ -factor covering all crops for our regression.

2.3 Correlation with Production Costs

As mentioned above, changes in yields levels are also related to changes in production costs, i.e. costs of all input factors used to produce one unit of output. We hypothesize that production costs per area have a functional relationship with yield level and agricultural land-use intensity. However, since the residuals in our data are not normally distributed, we cannot use a linear regression analysis. Instead, we have applied a correlation analysis between (a) yield and costs per area and (b) yield and costs per ton using the Pearson correlation coefficient (Rodgers and Nicewander, 1988) as well as the Kendall rank correlation coefficient (Kendall, 1938). We use two different correlation coefficients, to uncover potential, measure-related, biases in the analysis. Whereas the Pearson correlation coefficient measures the magnitude of the linear dependence between two variables, the Kendall rank correlation coefficient measures just any correlation based on a rank test (Kendall and Gibbons, 1990). Since residuals in our data are non-normally distributed, the significance of the Pearson test may be biased, if samples sizes are too small (Kowalski, 1972).

³ Infrastructure investments are composed of investments in transport, water and energy distribution, telecommunication and financial services, all related specifically to the agricultural sector according to GTAP 7.

Data for production costs are taken from the GTAP data base, version 7 (Narayanan and Walmsley, 2008), yield data are taken from FAOSTAT (2009). The data for small producing countries are less accurate and bring much noise into the analysis (Horridge and Laborde, 2008). Therefore, we decided only to take the top producing countries for each crop into account so that at least 90% of total crop production is included in the analysis. Moreover, we wanted to make sure that at least 1/3 of all available countries (31 countries) are included (an exception is oil palm, which is only produced in 20 countries worldwide).

2.4 Model Implementation

The global land-use model MAGPIE ("Model of Agricultural Production and its Impact on the Environment") has been developed to generate future land-use and agricultural production patterns, addressing a wide range of scenarios on population and income growth throughout the 21st century. It is a recursive dynamic model working on a regular spatial grid with a cell size of about $0.5^\circ \times 0.5^\circ$ (approximately $50 \times 50 km^2$ at the equator). The model works on ten-year time steps. On the biophysical side, it uses spatially explicit data on potential crop yields, land and water availability taken from the dynamic global vegetation model LPJmL (Bondeau et al., 2007). Economic data are used at the aggregate level of 10 economic world regions. For future demand trajectories the model derives specific land-use patterns and costs of agricultural production for each grid cell. These patterns are internally derived from external data for population (CIESIN et al., 2000) and gross domestic product (GDP) (World Bank, 2001) (see appendix A). The food energy demand for the year 1995 is taken from FAOSTAT (2008). The share of traded goods is kept constant over time and is based on self sufficiency ratios for the year 1995 (FAOSTAT, 2008). More information on model structure and features can be found in detail in Lotze-Campen et al. (2008, 2009); Popp et al. (2010). A mathematical description of the model is presented in appendix B.

Figure 2 shows a schematic overview of the endogenous implementation of technological change in MAGPIE. Investments in TC lead to a yield increase, which causes the τ -factor to rise. This implies an increase in production costs per area as well as a rise in the IY ratio. Hence, in order to achieve one unit of yield increase in a certain time step, a higher amount of TC investments has to be mobilized than in the previous period.

In addition, we have to consider some characteristics of the model and the agricultural sector. Typically, for endogenous technology implementations in economic models an intertemporal optimization approach is used due to the need of some kind of planning foresight (Ma and Nakamori, 2009). In contrast, MAGPIE is a recursive dynamic optimization model which solves each time step separately. To be able to reproduce planning foresight, in MAGPIE we use the annuity approach to transfer lump-sum TC investment to periodic payments including interest (Kellison, 1991). Another issue is the implementation of 15 years lag between R&D investment and yield impact. The model decides based on the expectations 15 years later how much should be invested. However, since there is no other cross-connection between these time steps, we are able to shift the investments to the time step when its impact takes place. This means: if the model needs yield growth in the year 2025 due to higher demand expectations, these 2025 model investments must have been made in 2010. However, the costs for R&D in 2010 in the model will be compounded and paid in 2025. This implementation allows for endogenizing technological change in a land-use model without using intertemporal optimization.

In order to validate our implementation we have compared long-term trends of simulated τ -factor development from 1995 to 2060 with observed data from 1960 to 2005, with a

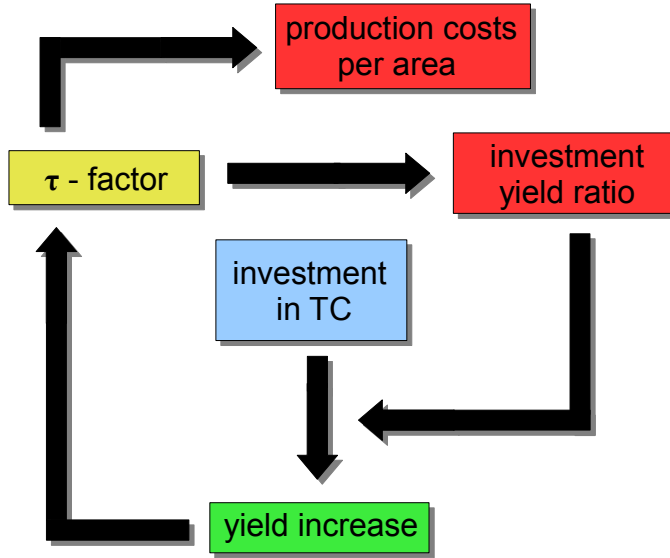


Fig. 2 Implementation of technological change in MAgPIE (schematic)

special focus on the overlap in 1995-2005. For the validation we took historical data from FAO on yield growth, which were neither part of our model parametrization nor calibration, and calculated the changes in τ -factor backwards starting from 2005.

3 Results

3.1 Regression and Correlation

The result of our regression between IY ratio and the τ -factor is the functional relationship of equation 3. Figure 3 shows the relationship in a graph for the 10 world regions of MAgPIE.

$$IY(\tau_i) = 9163 * \tau_i^{3.18} \quad (3)$$

The adjusted R-squared has a value of 0.94. F-test and t-tests are all significant at the 1% level. The elasticity between IY ratio and the τ -factor ϵ_{τ}^{IY} has the value of 3.18. As previously explained, changes in τ are proportional to changes in yield, and therefore we can transform this elasticity into an elasticity of yield with respect to accumulated TC investments (I), which is a more common representation (equation 4).

$$\epsilon_I^{yld} = \frac{1}{\epsilon_{\tau}^{IY} + 1} = 0.239 \quad (4)$$

Our result is close to the value of $\epsilon_I^{yld} = 0.296$, as reported by (Nelson et al., 2009).

With regard to the relationship between production costs and yield level, Table 2 shows the Pearson correlation coefficients and the Kendall rank correlation coefficients.

All correlations are positive and, except for oilcrops, they are all at least significant at the 10% level (most of them are significant at the 1% level). Table 3 shows the same information

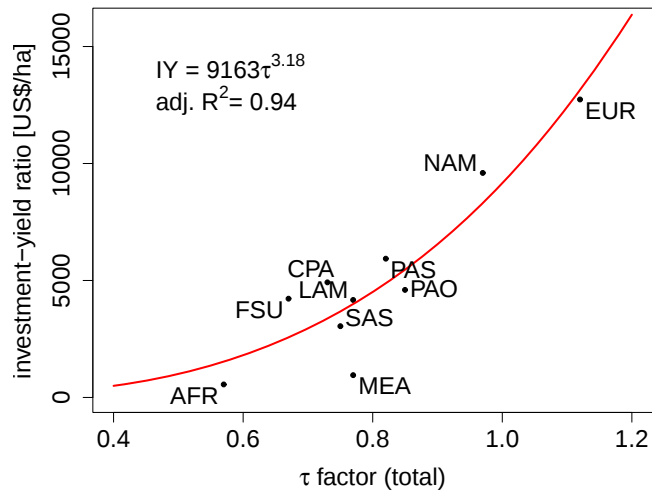


Fig. 3 investment-yield ratio in relation to τ -factor

crop types		Pearson		Kendall	
		correlation	p-value	correlation	p-value
cereals	temperate	0.36 **	0.05	0.24 *	0.06
	maize	0.45 ***	0.01	0.50 ***	0.00
	tropical	0.55 ***	0.00	0.50 ***	0.00
	rice	0.32 *	0.08	0.55 ***	0.00
oilcrops	fieldcrops	0.19	0.30	0.33 ***	0.01
	oil palm	0.18	0.44	0.08	0.63
sugar	cane	0.40 **	0.02	0.27 **	0.03
	beet	0.66 ***	0.00	0.55 ***	0.00
others	fibers	0.30 *	0.10	0.45 ***	0.00
	others	0.48 ***	0.01	0.60 ***	0.00

Table 2 Correlation between yield and production costs per area (* $p \geq 90\%$, ** $p \geq 95\%$, *** $p \geq 99\%$)

crop types		Pearson		Kendall	
		correlation	p-value	correlation	p-value
cereals	temperate	-0.21	0.25	-0.11	0.42
	maize	0.07	0.69	0.18	0.15
	tropical	-0.15	0.44	0.02	0.92
	rice	0.14	0.44	0.21 *	0.10
oilcrops	fieldcrops	0.01	0.95	-0.02	0.89
	oil palm	0.01	0.96	-0.08	0.63
sugar	cane	-0.06	0.76	0.05	0.74
	beet	0.35 *	0.06	0.28 **	0.03
others	fibers	0.05	0.80	0.19	0.14
	others	0.28	0.12	0.34 ***	0.01

Table 3 Correlation between yield and production costs per ton (* $p \geq 90\%$, ** $p \geq 95\%$, *** $p \geq 99\%$)

for the relationship between yields and production costs per ton. However, almost none of

the tested crop types shows a significant correlation. Comparing the results in both tables suggests the existence of a positive correlation between yields and area-related production costs, but no correlation between yield and output-related production costs. Based on this result we have implemented production costs per ton as a constant input for our model, which leads to a linear increase of production costs per area with yield.

crop types		costs [US\$/t]	used countries	share of total prod.
cereals	temperate	67	31	0.95
	maize	169	31	0.96
	tropical	158	31	0.99
	rice	182	31	0.99
oilcrops	fieldcrops	364	31	0.98
	oil palm	208	20	1.00
sugar	cane	15	31	0.99
	beet	38	31	0.99
others	fibers	731	31	0.98
	others	189	31	0.90

Table 4 Crop-specific, average costs per ton, number of countries used for averaging and the total production share of these countries

Table 4 shows the calculated costs per ton together with the number of countries included in this calculation and the share of total production covered by these countries. These costs per ton are used in MAGPIE for the calculation of production costs (see Appendix B).

3.2 Simulation Results

Figure 4 shows the development of the τ -factor (2005-2060) compared to past observations of the FAO (1960-2005). As an example we show the development path for maize since this is one of the most important crops and is grown in all parts of the world. It is taken as example since all other crop-types in our analysis show similar behaviour. Regions like Sub-Saharan Africa (AFR) and North America (NAM) show very strong increases in τ . However, the strongest increase is projected for the Middle East and North Africa region (MEA). This enormous increase is in line with FAO data for this region for the period since the 1980s. Overall three groups can be distinguished: Regions with increasing growth rates (MEA, AFR), constant rates (NAM, LAM, SAS and PAS) and decreasing rates (CPA, EUR, FSU, PAO), as compared to the past.

Figure 5 shows our model results compared with FAO observations in greater detail for the aggregate of all crops. It is important to note that the FAO data used for validation were not used as model input, neither as direct source, nor for calibration purposes. For direct comparing observations and model results, we focus on the overlap in 1995-2005. Moreover, the model results can be validated against the general trend in the observed data. Based on these criteria, the model results for EUR, LAM, MEA, NAM and SAS fit very well with observations. This means that they show the same long-term trends that have been noticed in the past (no bending and comparable long-term growth rates at the edge between observation and simulation) and that they also resemble observations between 1995 and 2005. Sub-Saharan Africa (AFR) shows significantly higher rates in our model compared to observations, especially in the long-term. FSU shows strong fluctuations in past data for reasons not captured by the model and is, therefore, hard to validate. PAS shows similar

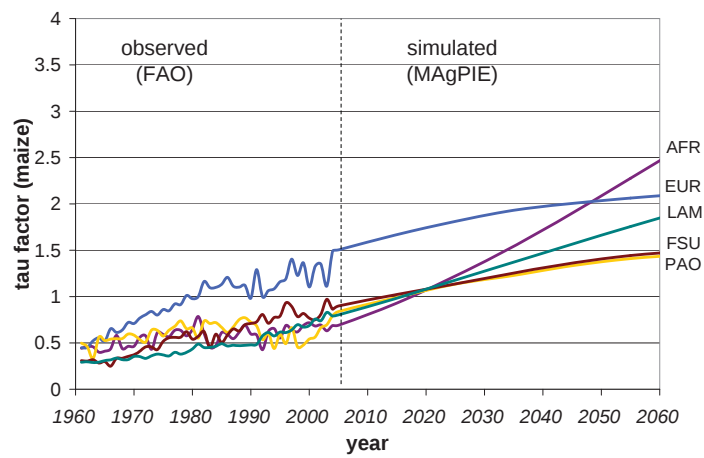
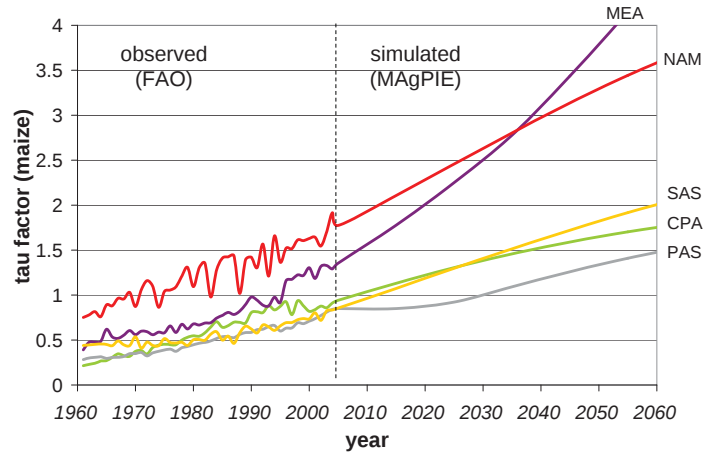


Fig. 4 Observed and simulated τ -factor for maize in the ten world regions

long-term trends in the future as compared to the past, while growth rates for PAO also seem to be slightly overestimated.

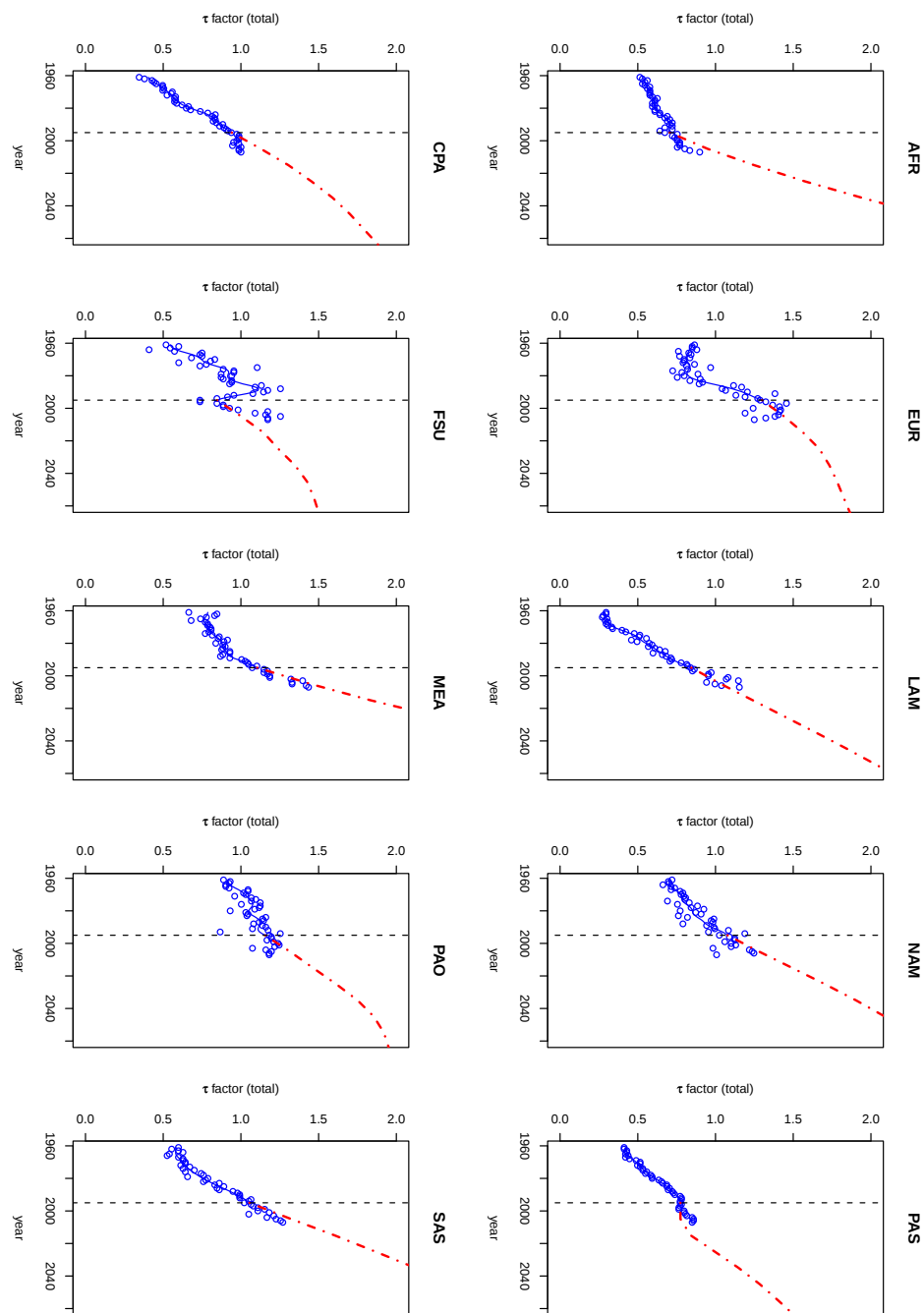


Fig. 5 Validation of MAGPIE model projections 1995-2060 (red chain line) with FAO observations 1960-2005 (blue dots) and its running mean (blue line)

4 Discussion

Technological change is a crucial driver for increasing agricultural yields. Our approach estimates the level and evolution of the investment-yield ratio relative to agricultural land-use intensity. The regression analysis confirms that a higher state of agricultural land-use intensity coincides with a higher IY ratio. Furthermore, our yield elasticity with respect to accumulated TC investments with the value 0.239 is in line with Nelson et al. (2009). Evenson (1989) showed that R&D spillovers are of major importance in agricultural research if the regions are small. However, increasing the region size, as in our implementation, reduces the role of spillover effects significantly. This means that our approximation is only valid at coarse scales and becomes invalid for more disaggregated regional units.

Results confirm that yields correlate with production costs per area. Since marginal production costs are constant, every additional production unit costs the same additional amount of money. Consequently, farmers will adopt the new technology since they expect higher yields at constant costs per ton.

Our validation of simulated output with observed data supports our model implementation. For six out of ten world regions we can reproduce the observed data (1995-2005) and the long-term trend very well. The stronger increase in AFR is due to its strong rise in projected agricultural demand. Since each region has to meet its internal demand in the model, AFR in the coming decades will have to invest in R&D a lot more than in the past. However, the observed data since 1995 already seem to show a switch to higher yield growth rates in Sub-Saharan Africa. For CPA and PAS we meet the long-term trend well, but the observed yield data since 1995 are stagnating, which we currently cannot explain. The countries of the former Soviet Union (FSU) are a special case, since the observed yield data in the past vary a lot, most likely due to the political transformation after 1990. Our simulation projects a fair rise in yields, which is in good agreement with the long-term trend, but contradictory to the observed, abrupt increase back to the 1990 yield level. Another reason for relatively weak validation results in some regions is that demand and trade are rather inflexible in the current version of MAGPIE and in some cases, like LAM or CPA, this might have strong impacts on future productivity levels. Notwithstanding, our overall validation results indicate the robustness of our approach, since the observed data are not considered in our analysis and are independent of the model results.

Our τ projections for maize provide rich insights with regard to future yield trends. The strong increase in Africa indicates what kind of yield growth rates be required to meet a soaring demand. North America, as the leading region for maize production, continues with high yield growth rates, but could be overtaken by the Middle East and North Africa region (MEA) in terms of growth rates. This region faces unfavourable cropping conditions and at the same time a higher demand increase and a strong political will to reduce food imports. If these conditions prevail in the future, huge investments in technological change would be required. In contrast, Europe continues along its trend over the past two decades when maize yields have not improved much. The Asian regions, starting from a lower yield level and facing a higher demand pressure in the future, have higher growth rates compared to Europe. Lastly, Latin America follows its strong yield growth path since the early 1990s, with high investments in the agricultural sector.

Our implementation of technological change can be used either in an intertemporal or a recursive dynamic optimisation approach. In our case, the latter option is favourable, as it reproduces the observed effect of continuous underinvestment in agricultural R&D (Ruttan, 1980; Roseboom, 2002). This market failure is caused by the limited foresight of decision makers concerning investments in R&D (Slaughter, 1996). An intertemporal optimization

model, however, would anticipate all the future benefits of R&D investments, which would lead to an optimal R&D investment path in R&D and an overestimation of yield increases, compared with observed trends.

5 Conclusion

During the lifetime of Thomas Malthus and before, growth in agricultural output was almost exclusively a result of growth in the use of input factors. This changed by the end of the 19th century and since then agricultural output has been mainly driven by increases in productivity. However, most agricultural sector and land-use models do not cover technological change as an endogenous driver. In order to fill this gap, we have presented a model approach for an endogenous implementation of technological change.

Our assumption, based on numerous studies, is that investments in technological change induce increases in land productivity. Our statistical analysis shows that the investment-yield ratio increases in a disproportionate way to the τ -factor and that production costs are linear correlated with the yield level. The results from the model MAgPIE indicate that regions with high demand projections, like Sub-Saharan Africa, or low potentials for land expansion, like Middle East and South Asia, have to make huge investments in future technological change. While the Middle East region and South Asia show this trend already in observed data, AFR shows this trend only since 1995. Hence, to meet its projected challenges in economic development and meet its growing agricultural demand, it seems indispensable for AFR to increase investments in R&D and infrastructure in order to achieve food security. Our endogenous implementation of technological change for reaching food security improves the projection quality of global agricultural models and is a further step towards producing more realistic future scenarios for agriculture.

References

- Alene, A. D., Coulibaly, O., 2009. The impact of agricultural research on productivity and poverty in sub-Saharan africa. *Food Policy* 34 (2), 198–209.
- Alston, J. M., 2000. A meta-analysis of rates of return to agricultural R&D: Ex pede Herculem? *Int Food Policy Res Inst IFPRI*.
- Alston, J. M., Beddow, J. M., Pardey, P. G., 2009. Agricultural research, productivity, and food prices in the long run. *Science* 325 (5945), 1209.
- Alston, J. M., Craig, B. J., Pardey, P. G., 1998. Dynamics in the Creation and Depreciation of Knowledge and the Returns to Research. *Environment and Production Technology Division, International Food Policy Research Institute*.
- Alston, J. M., Norton, G. W., Pardey, P. G., 1995. Science under scarcity: Principles and practice for agricultural research evaluation and priority setting. *Food systems and agrarian change (USA)*.
- Bondeau, A., Smith, P., Zaehle, S. O. N., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., Lotze-Campen, H., Müller, C., Reichstein, M., 2007. Modelling the role of agriculture for the 20th century global terrestrial carbon balance. *Global Change Biology* 13 (3), 679–706.
- CIESIN, IFPRI, WRI, 2000. Gridded population of the world (GPW), Version 2. Center for International Earth Science Information Network (CIESIN) Columbia University, Inter-

- national Food Policy Research Institute (IFPRI) and World Resources Institute (WRI), Palisades, New York.
- Cox, T., Mullen, J., Hu, W., 1997. Nonparametric measures of the impact of public research expenditures on Australian broadacre agriculture. *The Australian Journal of Agricultural and Resource Economics* 41 (3), 333–360.
- Dietrich, J. P., Schmitz, C., Müller, C., Fader, M., Lotze-Campen, H., Popp, A., 2010. Measuring agricultural land-use intensity. *Ecological Modelling* (submitted).
- Evenson, R. E., 1989. Spillover benefits of agricultural research: Evidence from U.S. experience. *American Journal of Agricultural Economics* 71 (2), 447–452.
- Evenson, R. E., Gollin, D., 2003. Assessing the impact of the green revolution, 1960 to 2000. *Science* 300 (5620), 758–762.
- Fan, S., Zhang, L., Zhang, X., 2002. Growth, inequality, and poverty in rural China: The role of public investments. *Int Food Policy Res Inst IFPRI*.
- FAOSTAT, 2008. Food & Agriculture Organization of the United Nations Statistics Division - <http://faostat.fao.org>, accessed 8/11/2008.
- FAOSTAT, 2009. Food & Agriculture Organization of the United Nations Statistics Division - <http://faostat.fao.org>, accessed 11/6/2009.
- Federico, G., 2005. *Feeding the world*. Princeton University Press.
- Heistermann, M., Müller, C., Ronneberger, K., Jun. 2006. Land in sight?: Achievements, deficits and potentials of continental to global scale land-use modeling. *Agriculture, Ecosystems & Environment* 114 (2-4), 141–158.
- Horridge, M., Laborde, D., 2008. TASTE a program to adapt detailed trade and tariff data to GTAP-related purposes. *GTAP Conference Paper* 2666.
- Huffman, W. E., Evenson, R. E., 2006. *Science for agriculture: A long-term perspective*. Wiley-Blackwell.
- Kellison, S., 1991. *The theory of interest*. McGraw-Hill/Irwin, 2nd edition.
- Kendall, M., Gibbons, J., 1990. *Rank Correlation Methods*. 5th e. London: Griffin.
- Kendall, M. G., 1938. A new measure of rank correlation. *Biometrika* 30 (1/2).
- Kowalski, C., 1972. On the effects of non-normality on the distribution of the sample product-moment correlation coefficient. *Applied Statistics* 20, 1–12.
- Lotze-Campen, H., Müller, C., Bondeau, A., Rost, S., Popp, A., Lucht, W., 2008. Global food demand, productivity growth and the scarcity of land and water resources: a spatially explicit mathematical programming approach. *Agricultural Economics* 39 (3), 325–338.
- Lotze-Campen, H., Popp, A., Beringer, T., Müller, C., Bondeau, A., Rost, S., Lucht, W., 2009. Scenarios of global bioenergy production: The trade-offs between agricultural expansion, intensification and trade. *Ecological Modelling* 221, 2188–2196.
- Ma, T., Nakamori, Y., Jul. 2009. Modeling technological change in energy systems – from optimization to agent-based modeling. *Energy* 34 (7), 873–879.
- Malthus, T. R., 1798. *An essay on the principle of population*. J. Johnson, London.
- Narayanan, B., Walmsley, T. L., 2008. *Global Trade, Assistance, and Production: The GTAP 7 Data Base*. Center for Global Trade Analysis, Purdue University.
- Nelson, G. C., Rosegrant, M. W., Koo, J., Robertson, R., Sulser, T., Zhu, T., Ringler, C., Msangi, S., Palazzo, A., Batka, M., Magalhaes, M., Valmonte-Santos, R., Ewing, M., Lee, D., 2009. *Climate change - impact on agriculture and costs of adaptation*. IFPRI Food Policy Report 21.
- Pardey, P. G., Beintema, N., 2001. *Slow Magic - Agricultural R&D a Century After Mendel*. Food Policy Report: International Food Policy Research Institute.
- Pardey, P. G., Beintema, N., Dehmer, S., Wood, S., 2006. *Agricultural Research a Growing Global Divide? Agricultural Science and Technology Indicators Initiative: International*

- Food Policy Research Institute.
- Pardey, P. G., Craig, B., 1989. Causal relationships between public sector agricultural research expenditures and output. *American Journal of Agricultural Economics*, 9–19.
- Popp, A., Lotze-Campen, H., Bodirsky, B., 2010. Food consumption, diet shifts and associated non-CO₂ greenhouse gases from agricultural production. *Global Environmental Change* 20, 451–462.
- Rodgers, J. L., Nicewander, W. A., Feb. 1988. Thirteen ways to look at the correlation coefficient. *The American Statistician* 42 (1), 59–66.
- Romer, P. M., 1990. Endogenous technological change. *The Journal of Political Economy* 98 (5), 71–102.
- Roseboom, J., 2002. Underinvestment in agricultural R&D revisited. *Quarterly Journal of International Agriculture* 41 (4), 297–316.
- Rosenzweig, M. R., Binswanger, H. P., McIntire, J., 1988. From land abundance to land scarcity - The effects of population growth on production relations in agrarian economies. Clarendon Press.
- Runge, C. F., Senauer, B., Pardey, P. G., Rosegrant, M., 2003. Ending Hunger in our Lifetime - Food Security and Globalization. The John Hopkins University Press.
- Ruttan, V. W., 1980. Bureaucratic productivity: the case of agricultural research. *Public Choice* 35 (5), 529–547.
- Schneider, U. A., Schwab, D. E., 2005. The european forest and agricultural sector optimization model. In: *Proceedings, IATRC Annual Meeting*, December. pp. 4–6.
- Slaughter, R. A., 1996. Long-term thinking and the politics of reconceptualization. *Futures* 28 (1), 75–86.
- Thirtle, C., Lin, L., Piesse, J., 2003. The impact of research-led agricultural productivity growth on poverty reduction in africa, asia and latin america. *World Development* 31 (12), 1959–1975.
- Trewavas, A., 2002. Malthus foiled again and again. *Nature* 418 (6898), 668–670.
- UN, 2005. World population prospects, the 2004 revision. UN Department of Economic and Social Affairs, New York, USA, URL: <http://www.un.org/esa/population/publications/sixbillion/sixbilpart1.pdf/>.
- van Meijl, H., van Tongeren, F., 1999. Endogenous international technology spillovers and biased technical change in agriculture. *Economic Systems Research* 11, 1.
- Verburg, R., Stehfest, E., Woltjer, G., Eickhout, B., 2009. The effect of agricultural trade liberalisation on land-use related greenhouse gas emissions. *Global Environmental Change* 19, 434–446.
- Wik, M., Pingali, P., Broca, S., 2008. Global agricultural performance: past trends and future prospects. Background paper for the World Development Report 2008.
- World Bank, 2001. World development indicators (CD-ROM), Washington DC.

A Population and GDP

year	1995	2005	2015	2025	2035	2045	2055	2065
AFR	552.67	743.16	926.07	1,125.00	1,312.86	1,481.03	1,629.10	1,753.19
CPA	1,280.64	1,480.41	1,581.95	1,650.63	1,673.05	1,677.04	1,658.75	1,632.24
EUR	554.44	589.45	586.15	574.98	559.02	532.27	505.17	479.74
FSU	276.34	292.75	295.00	295.00	285.36	275.14	261.64	245.94
LAM	452.00	550.36	622.55	686.75	738.79	779.81	809.56	830.30
MEA	277.74	357.18	423.05	485.67	541.49	589.73	632.76	670.58
NAM	292.11	331.96	354.87	375.03	390.91	399.85	403.78	403.13
PAO	133.78	145.75	147.79	146.72	145.72	144.47	139.94	131.68
PAS	383.23	462.27	516.86	565.23	614.21	652.18	674.22	684.05
SAS	1,269.92	1,572.17	1,796.92	1,997.73	2,148.85	2,265.16	2,347.33	2,397.64

Table 5 Population in million from 1995 to 2065 aggregated to ten world regions (CIESIN et al., 2000)

year	1995	2005	2015	2025	2035	2045	2055	2065
AFR	1,513	1,627	1,826	2,080	2,448	3,221	4,242	5,430
CPA	3,299	5,855	8,908	12,311	16,270	20,512	24,720	28,579
EUR	16,128	20,123	25,189	30,654	36,115	41,080	45,851	50,672
FSU	3,521	4,081	6,094	8,496	11,143	15,264	20,235	25,698
LAM	6,527	7,840	9,769	11,853	14,131	17,144	20,809	24,989
MEA	4,940	5,855	7,352	9,215	11,408	14,142	17,346	21,002
NAM	26,765	33,920	39,349	44,489	49,842	55,597	61,383	67,106
PAO	21,469	24,240	28,672	34,841	41,224	45,297	49,037	52,935
PAS	3,649	4,614	6,692	9,324	12,371	16,211	20,322	2,569
SAS	1,461	2,139	3,181	4,406	5,805	7,769	9,827	11,923

Table 6 GDP per capita (US\$ per number of people in purchasing power parities (PPP)) (World Bank, 2001)

B MAgPIE mathematical description

MAgPIE (Model of Agricultural Production and its Impact on the Environment) is a nonlinear recursive dynamic optimization model that links regional economic information with grid-based biophysical constraints simulated by the dynamic vegetation model LPJmL. A simulation run with the simulation period T can be described as a set

$$X = \{x_t \mid t \in T\} \subseteq \Omega \quad (5)$$

of solutions of a time depending minimization problem, i.e. for every timestep $t \in T$ the following constraint is fulfilled

$$\forall y \in \Omega : g_t(x_t) \leq g_t(y), \quad (6)$$

where the goal function for $t \in T$

$$g_t(x_t) = g(t, x_t, x_{t-1}, \dots, x_1, P_t) \quad (7)$$

depends on the solutions of the previous time steps $x_{t-1}, \dots, x_1$ and a set of time depending parameters P_t . We may interpret a MAgPIE simulation run $X = \{x_t \mid t \in T\} \subseteq \Omega$ as an element of the vector space $\Omega_T = \Omega \times T$.

B.1 Sets

The dimension of the domain Ω , on which for each timestep the minimization problem is defined, and of Ω_T depends on the following sets:

- $T = \{\text{time steps } t\}$: Simulation time steps, where t denotes the current time step, $t - 1$ the previous time step and so on. The first simulated time step is $t = 1$.
- $I = \{\text{world regions } i\}$: Economic world regions in MAgPIE.
- $J = \{\text{spatial cells } j\}$: Highest disaggregation level in MAgPIE.
- $K = \{\text{simulated products } k\}$: Union of vegetal products V and livestock products L ($K = V \cup L$).
- $L = \{\text{simulated livestock products } l\}$: Products simulated within the livestock sector of MAgPIE.
- $V = \{\text{vegetal products } v\}$: Products simulated within the crop sector of MAgPIE.
- $W = \{\text{water supply types } w\}$: Currently two types are implemented: rainfed 'rf' and irrigation 'ir'
- $C = \{\text{crop rotation groups } c\}$: Groups of crops, which produce similar effects in terms of crop rotation.

To highlight the substance of our model equations with regard to the agricultural and economic contents, we split our variable x_t into

$$x_t = \left(x_t^{area} \in \Omega^{area}, x_t^{prod} \in \Omega^{prod}, x_t^{tc} \in \Omega^{tc} \right) \in \Omega, \quad (8)$$

where the respective domains can be identified as the following vector spaces

$$\Omega^{area} = \mathbb{R}^{|J|} \times \mathbb{R}^{|V|} \times \mathbb{R}^{|W|} \quad (9)$$

$$\Omega^{prod} = \mathbb{R}^{|J|} \times \mathbb{R}^{|L|} \quad (10)$$

$$\Omega^{tc} = \mathbb{R}^{|I|} \quad (11)$$

As a result, we may specify the dimension of the solution space for each timestep as $\dim \Omega = |J| \cdot |V| \cdot |W| + |J| \cdot |L| + |I|$ and the dimension of $\Omega_T = \Omega \times T$ as $\dim \Omega_T = |T| \cdot \dim \Omega = |T| \cdot (|J| \cdot |V| \cdot |W| + |J| \cdot |L| + |I|)$.

In the following, variables and parameters are provided with subscripts to indicate the dimension of the respective subdomains. Subscripts written in quotes are single elements of a set. The order of subscripts in the variable, parameter and function definitions does not change. The names of variables and parameters are written as superscript.

B.2 Variables

Since MAGPIE is a recursive dynamic optimization model, all variables refer to a certain time step $t \in T$. In each optimization step, only the variables belonging to the current time step are free variables. For all previous time steps, values were fixed in earlier optimization steps. As we have seen above, we currently distinguish three variables $x_t^{area} \in \Omega^{area}$, $x_t^{prod} \in \Omega^{prod}$ and $x_t^{tc} \in \Omega^{tc}$ that can be described as follows:

- $x_{t,j,v,w}^{area}$: The total area of each vegetal production activity v for each water supply type w , each cell j and each time step t [ha]
- $x_{t,j,l}^{prod}$: The total production of each livestock product l , for each cell j at each time step t [ton dry matter]
- $x_{t,i}^{tc}$: The amount of yield growth triggered by investments in R&D [1]

B.3 Parameters

Besides variables, the model is fed with a set of parameters P . These parameters are computed exogenously and are in contrast to variables of previous time steps fully independent of any simulation output. Although most parameters are time independent, there exist also some parameters which are time dependent.

- $p_{t,j,v,w}^{yield}$: Yield potentials for each time step, each cell, each crop and each water supply type taking only natural variations into account and excluding changes due to technological change [ton/ha]
- $p_{t,i,k}^{dem}$: Regional food and material demand in each time step for each product [10^6 ton]
- $p_{i,l,k}^{fshr}$: Feed share describing the share of each product k of total feed production for livestock product l and corresponding transformation from GJ feed in ton dry matter [ton/GJ]
- $p_{i,l}^{feed}$: Feed requirements for each livestock product l in each region i [GJ/ton]
- $p_{i,k,l}^{byprod}$: Feed energy delivered by the byproducts of k that are available as feedstock for the livestock product l [GJ/ton]
- $p_{i,v}^{frv}$: Area related factor requirements for each crop and each region. The parameter is the product of observed yields in 1995 (FAOSTAT, 2009) and the production costs shown in table 4 [US\$/ha]
- $p_{i,l}^{frl}$: Production related factor requirements for livestock products for each livestock type and each region [US\$/ton]
- p_i^{lcc} : Area related land conversion costs for each region [US\$/ha]
- p_i^{tcc} : Technological change costs factor containing an interest correction, an expected lifetime factor and a general cost factor [US\$/ha]
- $p_{i,v}^{\tau 1}$: τ -Factor representing the agricultural land use intensity in the first simulation time step for each crop in each region [1]
- p^{exp} : Correlation Exponent between τ -Factor and technological change costs [1]
- $p_{i,v}^{seed}$: Share of production that is used as seed for the next period calculated for each crop in each region [1]
- $p_{t,i,k}^{xs}$: Regional excess supply for each product and each time step describing the amount produced for export [10^6 ton]
- $p_{i,k}^{sf}$: Regional self sufficiencies for each product [1]
- p^{tb} : Trade balance reduction factor. This factor is always less or equal 1 and is used to relax the trade balance constraints depending on the particular trade scenario.
- p_j^{land} : Total amount of land available for crop production in each cell [10^6 ha]
- $p_j^{ir.land}$: Total amount of land equipped for irrigation in each cell [10^6 ha]
- $p_{j,k}^{watreq}$: Cellular water requirements for each product [$m^3/ton/a$]
- p_j^{water} : Amount of water available for production in each cell [$m^3/ton/a$]
- p_c^{rmax} : Maximum share of crop groups in relation to total agricultural area [1]
- p_c^{rmin} : Minimum share of crop groups in relation to total agricultural area [1]

[all ton units in dry matter]

B.4 Sub-functions

To lighten the general model structure, some model components which appear more than once in the model description and depend on the variables of the current time step t are arranged as functions:

$$f_{t,i}^{growth}(x_t) = \prod_{\tau=1}^t (1 + x_{\tau,i}^{tc}) \quad (12)$$

$$f_{t,i,k}^{prod}(x_t) = \sum_{j_i} \begin{cases} x_{t,j,k}^{prod} & : k \in L \\ \sum_w x_{t,j,k,w}^{area} p_{t,j,k,w}^{yield} f_{t,i}^{growth}(x_t) & : k \in V \end{cases} \quad (13)$$

$$f_{t,i,k}^{dem}(x_t) = p_{t,i,k}^{dem} + \sum_l p_{i,l,k}^{fshr} \left(p_{i,l}^{feed} f_{t,i,l}^{prod}(x_t) - \sum_{\kappa} p_{i,\kappa,l}^{byprod} f_{t,i,\kappa}^{prod}(x_t) \right). \quad (14)$$

- $f_{t,i}^{growth}$: Growth function describing the aggregated yield amplification due to technological change compared to the level in the starting year for each year t and region i .
- $f_{t,i,k}^{prod}$: Function representing the total regional production of a product k in region i at timestep t . In the case of vegetal products, it is derived by multiplying the current yield level with the total area used to produce this product. In the case of livestock products, it is represented by the related production variable.
- $f_{t,i,k}^{dem}$: Function defining the demand for product k in region i at timestep t . It consists of an exogenous demand for food and materials $p_{t,i,k}^{dem}$ and an endogenous demand for feed, which is calculated as the feed demand generated by the livestock production minus the feed supply gained through byproducts.

B.5 Goal function

$$g_t(x_t) = g(t, x_t, x_{(t-1)}, \dots, x_1, P_t) \quad (15)$$

The goal function describes the value that is minimized in our recursive dynamic optimization model structure in each timestep. It is time dependent, i.e it differs for each time step, depending on the solutions of the previous time steps. We define the goal function as follows:

$$\begin{aligned} g_t(x_t) = & \sum_{i,v} \left(p_{i,v}^{frv} f_{t,i}^{growth}(x_t) \sum_{j_i,w} x_{t,j,v,w}^{area} \right) \\ & + \sum_{i,l} \left(p_{i,l}^{frl} f_{t,i,l}^{prod}(x_t) \right) \\ & + \sum_i \left(p_i^{lcc} \sum_{j_i,v,w} (x_{t,j,v,w}^{area} - x_{t-1,j,v,w}^{area}) \right) \\ & + p^{tcc} \sum_i \left(x_{t,i}^{tc} \left(\frac{1}{|V|} \sum_v p_{i,v}^{\tau 1} f_{t,i}^{growth}(x_t) \right)^{p^{exp}} \sum_{j_i,v,w} x_{t-1,j,v,w}^{area} \right). \end{aligned} \quad (16)$$

The function describes the total costs of agricultural production. The total costs can be splitted in four terms: 1. The area depending factor costs of vegetal production, which increase with the yield gain due to technological development. 2. The factor costs of livestock production depending on the production output. 3. The land conversion costs which arise, when non-agricultural land is cleared and prepared for agricultural production. 4. The costs, which arise by investing in technological development to increase yields by new inventions and improvements in management strategies. The technological change costs are proportional to the total cropland area of a region and increase disproportionate with the yield growth bought in the current timestep and the agricultural land-use intensity.

B.6 Constraints

Constraints are used to describe the boundary conditions, under which the goal function is minimized.

B.6.1 Global demand constraints (for each activity k)

$$\sum_i \frac{f_{t,i,k}^{prod}(x_t)}{1 + p_{i,k}^{seed}} \geq \sum_i f_{t,i,k}^{dem}(x_t) \quad (17)$$

These constraints describe the global demand for agricultural commodities: The total production of a commodity k adjusted by the seed share required for the next production iteration has to meet the demand for this product.

B.6.2 Tradebalance (for each region i and product k)

$$\frac{f_{t,i,k}^{prod}(x_t)}{1 + p_{i,k}^{seed}} \geq p^{tb} \begin{cases} f_{t,i,k}^{dem}(x_t) + p_{t,i,k}^{xs} & : p_{i,k}^{sf} \geq 1 \\ f_{t,i,k}^{dem}(x_t) p_{i,k}^{sf} & : p_{i,k}^{sf} < 1 \end{cases} \quad (18)$$

The trade balance constraints are similar to the global demand constraints, except that it acts on a regional level. In the case of an exporting region (self sufficiency for the product k is greater than 1), the production has to meet the domestic demand supplemented by the demand caused due to export. In the case of importing regions (self sufficiency less than 1), the domestic demand is multiplied with the self sufficiency to describe the amount which has to be produced by the region itself. In both cases the demand is multiplied with a so called "trade balance reduction factor". This factor is always less or equal 1 and is used to relax the trade balance constraints depending on the particular trade scenario, that is run.

B.6.3 Land constraint (for each cell j)

$$\sum_{v,w} x_{t,j,v,w}^{area} \leq p_j^{land} \quad (19)$$

$$\sum_v x_{t,j,v,ir'}^{area} \leq p_j^{ir.land} \quad (20)$$

The land constraints guarantee, that no more land is used for production than available. The first set of land constraints ensures the land availability for agricultural production in general. The second one secures, that irrigated crop production is restricted to areas that are equipped for irrigation.

B.6.4 Water constraints (for each cell j)

$$\sum_v x_{t,j,v,ir'}^{area} p_{t,j,v,ir'}^{yield} f_{t,i(j)}^{growth}(x_t) p_{j,v}^{watreq} + \sum_l x_{t,j,l}^{prod} p_{j,l}^{watreq} \leq p_j^{water} \quad (21)$$

In MAGPIE, the production of animal commodities as well as vegetal goods produced with irrigation requires water. The required amount of water is proportional to the production volume. The whole cellular water demand must be less or equal to the water available for production in this cell.

B.6.5 Rotational constraints (for each crop rotation group c , cell j and irrigation type w)

$$\sum_{v_c} x_{t,j,v,w}^{area} \leq p_c^{rmax} \sum_v x_{t,j,v,w}^{area} \quad (22)$$

$$\sum_{v_c} x_{t,j,v,w}^{area} \geq p_c^{rmin} \sum_v x_{t,j,v,w}^{area} \quad (23)$$

The rotational constraints are used to describe crop rotations, but also other aspects such as cultural preferences or efforts of autonome food production systems. This is achieved by defining for each vegetal product a maximum and minimum share relative to total area under production in a cell. While crop rotation structures are exclusively described with the maximum share constraints, cultural preferences and autonomy efforts are basically described with the minimum constraints.