

Aggregation of spatial units in linear programming models to explore land use options

R.J. HIJMANS¹ AND M.K. VAN ITTERSUM*

Department of Theoretical Production Ecology, Wageningen Agricultural University, P.O. Box 430, NL-6700 AK Wageningen, The Netherlands

¹ Present address: International Potato Center (CIP), Apartado 1558, Lima 100, Peru

* Corresponding author (fax +31-317-484892; e-mail: martin.vanittersum@staff.type.wau.nl)

Received 21 July 1995; accepted 19 March 1996

Abstract

Quantitative information on land use can be confronted with social, economic and environmental objectives in interactive multiple-goal linear programming (IMGLP) models to explore land use options. For various reasons, quantitative information on land use (input-output combinations) and resources is often aggregated to administrative units when integrated in the IMGLP model. In this paper, consequences of aggregating spatial units in an IMGLP model are analyzed for both a schematized and an existing IMGLP model (*GOAL*) exploring land use options for the European Union. A discrimination was made between effects on objective functions for the system as a whole, and effects on related optimum land use allocation within the system. In *GOAL*, effects on land use allocation tend to be more important than effects on the value of objective functions. Several rules or factors were identified that determine the effect of aggregation, among which the degree in curvilinearity in input-output relations and the method of aggregation are important ones. However, because of many complicated interacting effects, the aggregation error is difficult to predict. Therefore, in land use studies using IMGLP, the motto is 'first optimize the linear programming model at the non-aggregated level and then aggregate to the appropriate policy level'. If aggregation is inevitable because LP models become too big, aggregation according to agro-ecological criteria, i.e., aggregation of units with similar output-input ratios and constraints, results in the smallest errors.

Keywords: linear programming, IMGLP, aggregation, land use allocation, land use options

Introduction

In studying future land use options, systems analysis and modelling are indispensable tools, because the systems under study are generally unique systems that do not allow experimentation (De Wit, 1993). A model is a simplified representation of a part of reality. In models of agro-ecosystems, reality is usually simplified in time, space and the number of processes considered per unit of time or space. It is impor-

tant to be aware of the limitations inherent to this abstraction process (O'Neill & Rust, 1979). In crop growth models, for example, atmospheric processes are reduced to only a few variables, such as radiation and temperature. In addition, this information is aggregated in time. The degree of aggregation of weather in time highly affects the results of simulations. Nonhebel (1994) showed that the use of monthly instead of daily weather data may cause deviations of 50% in the prediction of water-limited wheat yields. The use of crop growth simulation models requires an appropriate time scale and time steps.

In quantitative land evaluations and land use studies, the level of spatial aggregation in models is very important. The level of spatial detail should be dictated by the objectives of the study, the appropriate scale to present the results, and the heterogeneity of the land units within the system under investigation. It is, however, also determined by pragmatic arguments such as the level of detail of available soil maps, climate data and computing capacity. Consequences of the use of aggregated units are often not explicitly analyzed, although this may lead to considerable errors (cf. De Wit & Van Keulen, 1987). Much of the spatial variation at a detailed level is obscured when, for instance, land evaluation units are aggregated into larger administrative units. Sometimes the aggregation errors are difficult to trace and estimate. This may be particularly true in land use studies in which the results of a land evaluation and an analysis of agricultural production activities, often after aggregation of information, are used as input for a linear programming model exploring land use options (Rabbinge & Van Ittersum, 1994; De Koning *et al.*, 1995).

This paper analyzes the effects of using aggregated spatial units (usually administrative regions heterogeneous in soil type and climate) in interactive multiple-goal linear programming (IMGLP) models used to explore land use options (De Wit *et al.*, 1988; Veeneklaas *et al.*, 1991). First, the IMGLP procedure and the relevance and inevitability of aggregation in IMGLP models are briefly explained. Then some rules and factors that determine the effect of aggregating spatial units in IMGLP models are identified and illustrated with a schematic example of a linear programming model for land use. Finally, effects of aggregation in the *GOAL* model (General Optimal Allocation of Land use); an IMGLP model that has been used for exploring land use options in the European Community (Rabbinge & Van Latesteijn, 1992), are analyzed.

IMGLP and the relevance of aggregation

In various studies exploring land use options, the IMGLP method has proven to be a useful instrument (De Wit *et al.*, 1988; Van Keulen, 1990; Veeneklaas *et al.*, 1991; Rabbinge *et al.*, 1994). IMGLP is a procedure using the linear programming (LP) technique. It can be used to confront technical information on agricultural production possibilities with various, often conflicting, socio-economic and environmental objectives, enabling the exploration of land use options under various policy views in a region. This may help to make explicit the consequences of different aims for land use.

An IMGLP model consists of (1) quantitative descriptions of relevant land use activities (e.g. growing a crop on a specified unit of land with a certain production technique) in terms of inputs and outputs, (2) a set of technical or physical constraints related to these activities (e.g. restricting the use of finite resources such as land and water) and (3) objective functions distilled from the prevailing policy views in the region under study, which can either be optimized or used as normative constraints (e.g., minimization of the use of pesticides). In each run, one objective function is optimized while the other objectives are used as constraints. Upper or lower bounds can be put on these 'goal constraints'. For different policy views land use scenarios can be generated by optimizing and putting bounds on the most relevant combination of objectives. These land use scenarios have two types of results: the optimum values of the objective functions and the corresponding optimal land use.

For several reasons, information on agricultural land use based on a land evaluation is often available at a much more detailed scale than that used in the LP model. Policymakers are usually interested in figures for administrative units, such as municipalities, provinces or countries, and not for small land evaluation units. Moreover, the size of an IMGLP model is mainly determined by the number of land use activities and constraints specified per spatial unit and the number of spatial units. When defining land use activities for each of the small land evaluation units, the LP model may become huge and impossible to solve, even with sophisticated software and hardware. Therefore, land use activities and constraints for IMGLP models are usually defined for administrative units, each containing aggregated land evaluation units. It is evident from other studies (cf. Nonhebel, 1994) that the moment and way of aggregating information may have enormous effects on the results of modelling studies.

The effect of aggregating information in LP models has been investigated only for economic studies. Several authors concluded that whatever the aggregation method used in LP models, there will always be an aggregation error, except in some rather restricted theoretical cases (Day, 1963; Paris & Reuser 1973; Zipkin 1980a,b) with limited scope for practical application (Spren & Takayama, 1980). No studies on the effect of aggregating spatial information in LP models for land use are available.

Schematic example of aggregation in linear programming models

The LP model

The objective function of the schematic LP model is minimization of the use of input I (e.g., kg nitrogen) for production of a certain output O of a crop (e.g., kg wheat) in a small system. The crop can be grown in 6 homogeneous (in soil and climate) land evaluation units, a, b, c, d, e, f , of 1 ha each. For each of these spatial units, only one production system is considered. The fictive input-output data are presented in Figure 1a, with output O on the x-axis and input I on the y-axis (for quantification of input-output combinations of agricultural activities, the target-oriented approach is often used: the identification of the optimal combination of inputs to realize a partic-

ular output level). Table 1 presents the mathematical formulation of the LP model.

In the optimization, production is allocated to the spatial units with the highest output/input (O/I) ratio. Figure 1a represents these ratios as the inverse of the slope of the lines through the origin and each of the points, i.e., the ratio is highest in a , followed by b , etc.

The minimum input I required to produce a certain output O can be read from Figure 1b, representing cumulative input against cumulative output, with the spatial units ordered according to their output/input ratio. In the optimization, wheat production is first allocated to a , then, when all land in a has been used, to unit b , and so on, until the required output has been realized.

Both activities and constraints are usually defined per spatial unit. Consequently, aggregating spatial units implies that both activities and constraints are aggregated. Because aggregation of activities and of constraints may have different effects, they are discussed separately (cf. Egbert & Kim, 1975; Zipkin, 1980a,b).

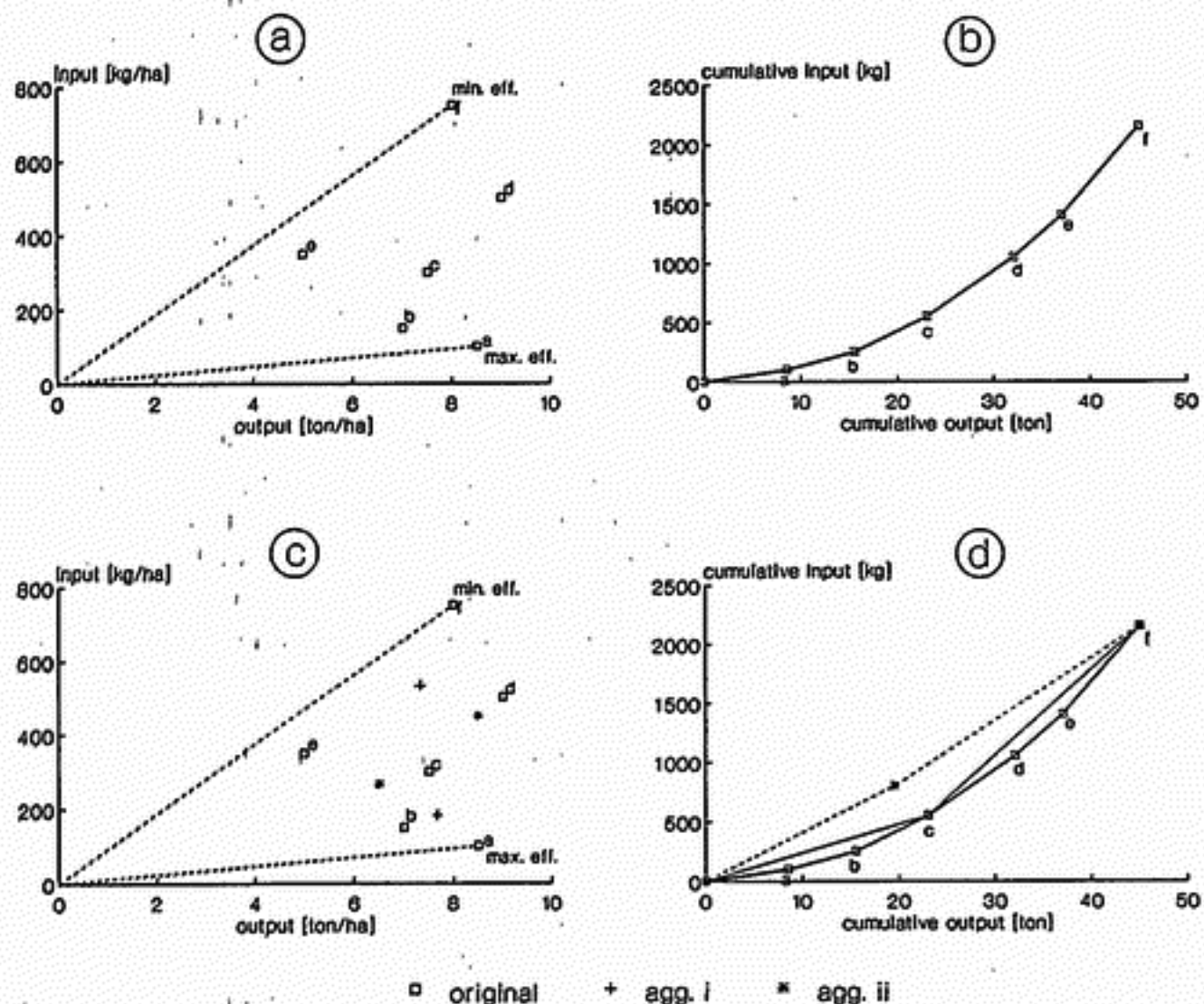


Figure 1. Fictive output-input relations for a fictive crop in regions a to f. (a) - non-aggregated output and input, and lines indicating maximal and minimal efficiency (output/input ratio). (b) - the non-aggregated cumulative output and input; regions have been ordered according to their output/input ratio. (c) - output and input before and after two ways of aggregation and lines indicating maximal and minimal efficiency (output/input ratio). (d) - cumulative output and input before and after two ways of aggregation; regions have been ordered according to their output/input ratio.

LINEAR PROGRAMMING MODELS TO EXPLORE LAND USE OPTIONS

Table 1. Model formulation of the schematic example before and after aggregation of the spatial units (activities) *a, b, c, d, e* and *f*. In aggregation i regions *a, b* and *c* and regions *d, e* and *f* were aggregated. See also Figure 1.

Before aggregation

Min $\{100x_a + 150x_b + 300x_c + 500x_d + 350x_e + 750x_f\}$ (objective function)

subject to

$8.5x_a + 7x_b + 7.5x_c + 9x_d + 5x_e + 8x_f = \text{Output}$ (production constraint)

$x_a, x_b, x_c, x_d, x_e, x_f \leq 1$ (area constraints)

After aggregation of activities a, b, c and d, e, f (aggregation i)

Min $\{183x_{abc} + 533x_{def}\}$ (objective function)

subject to

$7.67x_{abc} + 7.33x_{def} = \text{Output}$ (production constraint)

$x_{abc}, x_{def} \leq 3$ (area constraints)

* x_a : area under crop with input 100 kg/ha and output 8.5 tons per ha in region *a*.

Aggregation of spatial units: activities

In the aggregated model, only two spatial units are distinguished, each containing three of the non-aggregated land evaluation units. Inputs and outputs for activities in these new spatial units of 3 ha each are derived by averaging those for two groups of three non-aggregated units. Because there is only one constraint considered in this example, i.e. the area of the land units, and because the land units are all 1 ha, we can disregard the effect of aggregation of constraints here. Ten combinations of non-aggregated units are possible; but which combination should be selected?

Figure 1c presents two ways of aggregating: (i) according to the criterion of similar output/input ratio ($a+b+c, d+e+f$), and (ii) according to the criterion of similar output ($a+d+f, b+c+e$). Table 1 shows the mathematical formulation of the model after aggregation i. Both aggregations result in a reduction in the variation in output/input ratios among the spatial units. After aggregation i, this reduction is much smaller than after aggregation ii.

In Figure 1d, cumulative input is plotted against cumulative output before and after aggregation i and ii, with the spatial units again being ordered according to their output/input ratio. The difference between the non-aggregated and aggregated model in input required to produce a given output is defined as the aggregation error \mathcal{A} . In the optimization, the units are selected in order of their output/input ratio; the regions with the highest output/input ratio are selected first. Therefore, the aggregation in which units with similar output/input ratios are aggregated (aggregation i) results in the smallest aggregation error \mathcal{A} . For both aggregation procedures, \mathcal{A} is non-negative (when maximizing an objective function, \mathcal{A} is non-positive), and \mathcal{A} depends on the level of required output or, in other words, the area required to produce that output. For any similarly weighted aggregation, $\mathcal{A} = 0$ when the required level of production is such that the total area (6 ha) must be used, or no area is used. In aggregation i, it also holds that $\mathcal{A} = 0$ when 3 ha is used, because then *a, b* and *c* are selected in both the non-aggregated and aggregated model.

Aggregation of spatial units: constraints

So far, only an area constraint has been considered in this example: in each unit, no more land can be used than is available for crop production. We now introduce an additional constraint: the availability of irrigation water in each unit. We assume that in units *a*, *c* and *e* in fact 2 ha are suitable for crop production, but because of the lack of irrigation water, only 1 ha can be used, whereas in units *b*, *d* and *f*, of 1 ha each, surplus water is available (Table 2). Furthermore, the same input-output data are used as in the example of Table 1.

The simplest way to aggregate the water constraints for the various units is by adding pools of individual units. In physical terms, this implies that it is possible to transport water among land evaluation units. The total amount of water is not affected by this aggregation, but the distribution changes. When aggregating the water constraints for units *a-c* and *d-f* (aggregation i), it is possible to irrigate 1 ha extra in unit *a* and 1 ha extra in unit *e*. This results in an increase in maximum production of $8.5 + 5 = 13.5$ tons (Figure 2a). The input required to obtain a certain output decreases and the maximum production in the total system increases (Figure 2a) because, after aggregation of water constraints, water can be applied in regions where it is used most efficiently.

Contrary to the aggregation of activities, it is impossible to formulate a general rule for the best aggregation of constraints. Practical, agro-ecological and economic implications of aggregating constraints should be taken into account, in this case whether the transport of water is practically and economically feasible. Instead of aggregating constraints by addition, different aggregation rules could be used. In this example, we could use a (weighted) average of the individual water pools, or, to be on the safe side, take the water pool of the smallest unit of an aggregate. It de-

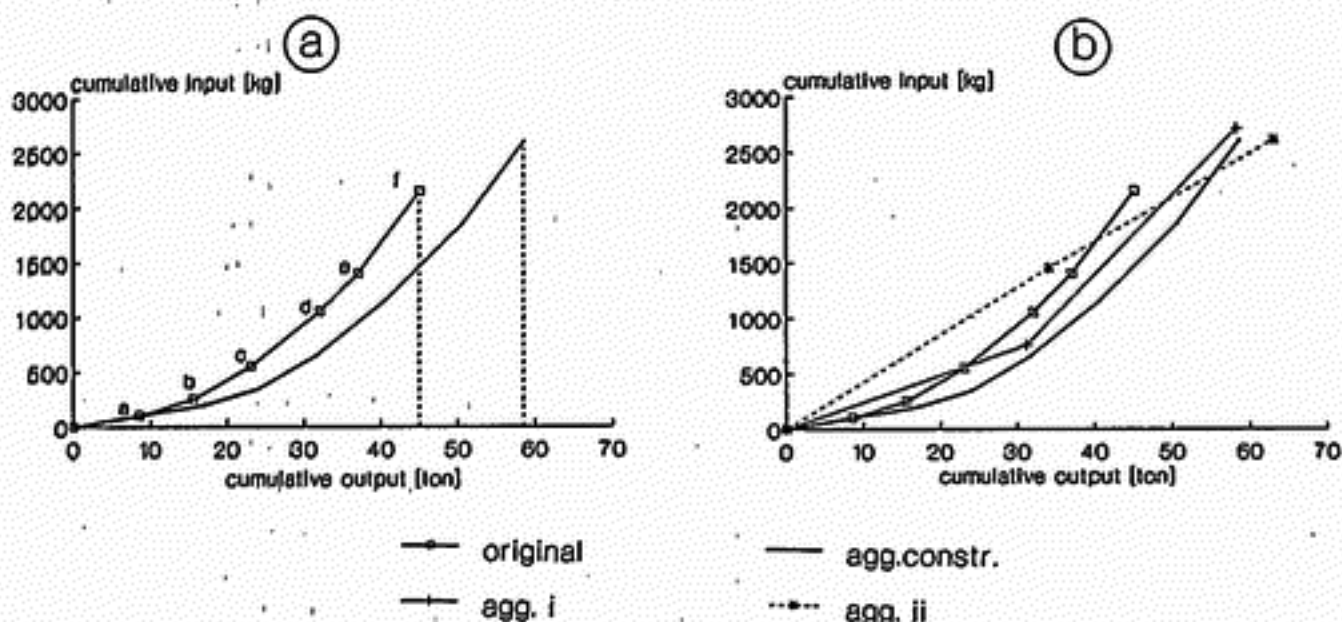


Figure 2. (a) – the cumulative output-input relations before and after aggregation of water constraints; regions have been ordered according to their output/input ratio. (b) – cumulative output-input relations before and after two ways of aggregation of constraints and activities.

LINEAR PROGRAMMING MODELS TO EXPLORE LAND USE OPTIONS

Table 2. Model formulation of the schematic example before and after aggregation of the simplified water constraints for units *a, b, c, d, e* and *f*. In aggregation i the water constraints for regions *a, b* and *c* and regions *d, e* and *f* were aggregated. See also Figure 2.

Before aggregation

Min $\{100x_a + 150x_b + 300x_c + 500x_d + 350x_e + 750x_f\}$ (objective function)

subject to

$8.5x_a + 7x_b + 7.5x_c + 9x_d + 5x_e + 8x_f = \text{Output}$ (production constraint)

$x_a, x_c, x_e \leq 2; x_b, x_d, x_f \leq 1$ (area constraints)

$x_a, x_c, x_e \leq 1; x_b, x_d, x_f \leq 2$ (simplified water constraints)

After aggregation of water constraints for units a, b, c and d, e, f (aggregation i)

Min $\{100x_a + 150x_b + 300x_c + 500x_d + 350x_e + 750x_f\}$ (objective function)

subject to

$8.5x_a + 7x_b + 7.5x_c + 9x_d + 5x_e + 8x_f = \text{Output}$ (production constraint)

$x_a, x_c, x_e \leq 2; x_b, x_d, x_f \leq 1$ (area constraints)

$x_a + x_b + x_c \leq 4; x_d + x_e + x_f \leq 5$ (simplified water constraints)

* x_a : area under crop with input 100 kg/ha and output 8.5 tons per ha in region *a*.

depends on the specific system and the objectives of the study, how this problem should be approached, and to what aggregation error it will lead.

Aggregation of spatial units: activities and constraints

Often, the consequence of aggregating spatial units is that activities *and* constraints are aggregated. It is impossible not to aggregate constraints when aggregating activities (e.g., when crop activities are defined at the aggregated level, water constraints for these activities cannot be quantified at the non-aggregated level). Therefore, the effect of aggregating spatial units in LP models is usually a combined effect of the aggregation of activities and of constraints. Figure 2b shows how \bar{A} depends on both the aggregation method and the required total output (or the area required to produce that output). Both aggregation i and ii may lead to a higher and lower use of I , depending on O . After both aggregations i and ii, a higher maximum O can be realized than without any aggregation: the maximum production for the total system increases because water can be used more efficiently.

To conclude, minimizing an objective function results in $\bar{A} \geq 0$ (a higher minimum solution), when activities of spatial units are aggregated by averaging the input-output coefficients, or in $\bar{A} \leq 0$ (a lower minimum solution) when aggregating resource use constraints of spatial units by addition. Analogously, it can be shown that the opposite holds when maximizing an objective function: then $\bar{A} \leq 0$ when aggregating activities or $\bar{A} \geq 0$ when aggregating constraints. In fact, aggregation of activities reduces the feasible space whereas aggregation of constraints enlarges it. The effect of aggregating activities *and* constraints is a combination of these effects. The effect of aggregation on the optimum value of the objective function can be substantial but it depends on the aggregation method and the area used for agricultural production (see also Discussion and conclusions).

Aggregation in *GOAL*

Description of GOAL

The effects of aggregating spatial units were also investigated in a complete IMGLP land use model, that is a simplified version of *GOAL*, a model for the exploration of land use options in the European Union (Schéele, 1992; Rabbinge & Van Latesteijn, 1992; Rabbinge *et al.*, 1994). The simplified *GOAL* version has the same basic structure as the original model. The only essential differences are that forestry and perennial cropping were not included and animal husbandry activities and water constraints were simplified (Van Ittersum *et al.*, 1995).

The activities in *GOAL* include crop, roughage and livestock production and industrial conversions of agricultural products. The constraints include balances: no lack or surplus is allowed for food and feed and no surplus for manure. Other constraints set upper limits to the use of resources: area constraints and water constraints. Most of the activities and constraints are defined per spatial unit. Different levels of demand for agricultural products (levels of required production) at the European Union level are distinguished, depending on the human diet and trade situation. In the present study, demand for agricultural products is based on the current diet and on autarky in the EU. *GOAL* contains agro-technical, socio-economic and environmental objective functions that were defined at the EU level: the minimization of agricultural area and costs of agricultural production, maximization of agricultural employment and minimization of nitrogen loss and use of pesticides. Thus, in the *GOAL* model, resource use or loss can be minimized (land, costs, nitrogen and pesticides) or maximized (labour) for a given agricultural production.

Technical coefficients for land use activities were derived from the results of a mixed qualitative/quantitative land evaluation, using a geographical information system in combination with crop growth simulation models (Van Lanen *et al.*, 1992; De Koning & Van Diepen, 1992; De Koning *et al.*, 1995). Soil characteristics were derived from the digitized soil map of the EU (1:1,000,000). Climate data from 109 meteorological stations were used. The overlay of the soil map, the climatic map and a map with the administrative regions (NUTS-1 regions) resulted in ca. 4200 different land evaluation units. For each of these units, input-output coefficients for various production activities were generated. This resulted in about 120 land use activities per land evaluation unit and, given the 4200 land evaluation units, about 500,000 activities for the EU in total. Policymakers are primarily interested in results at the administrative levels, such as the NUTS-1 level, and most statistical data are available at this NUTS-1 level (De Koning *et al.*, 1995). In addition, there are serious practical (software, time) limitations to working with a linear programming model containing 500,000 activities (De Koning *et al.*, 1995). Therefore, the 4200 land evaluation units with their input-output combinations and constraints were aggregated to 58 NUTS-1 regions (Fig. 3a). The LP model at the NUTS-1 level was the starting point for investigating the effects of aggregating spatial units in *GOAL*.

LINEAR PROGRAMMING MODELS TO EXPLORE LAND USE OPTIONS

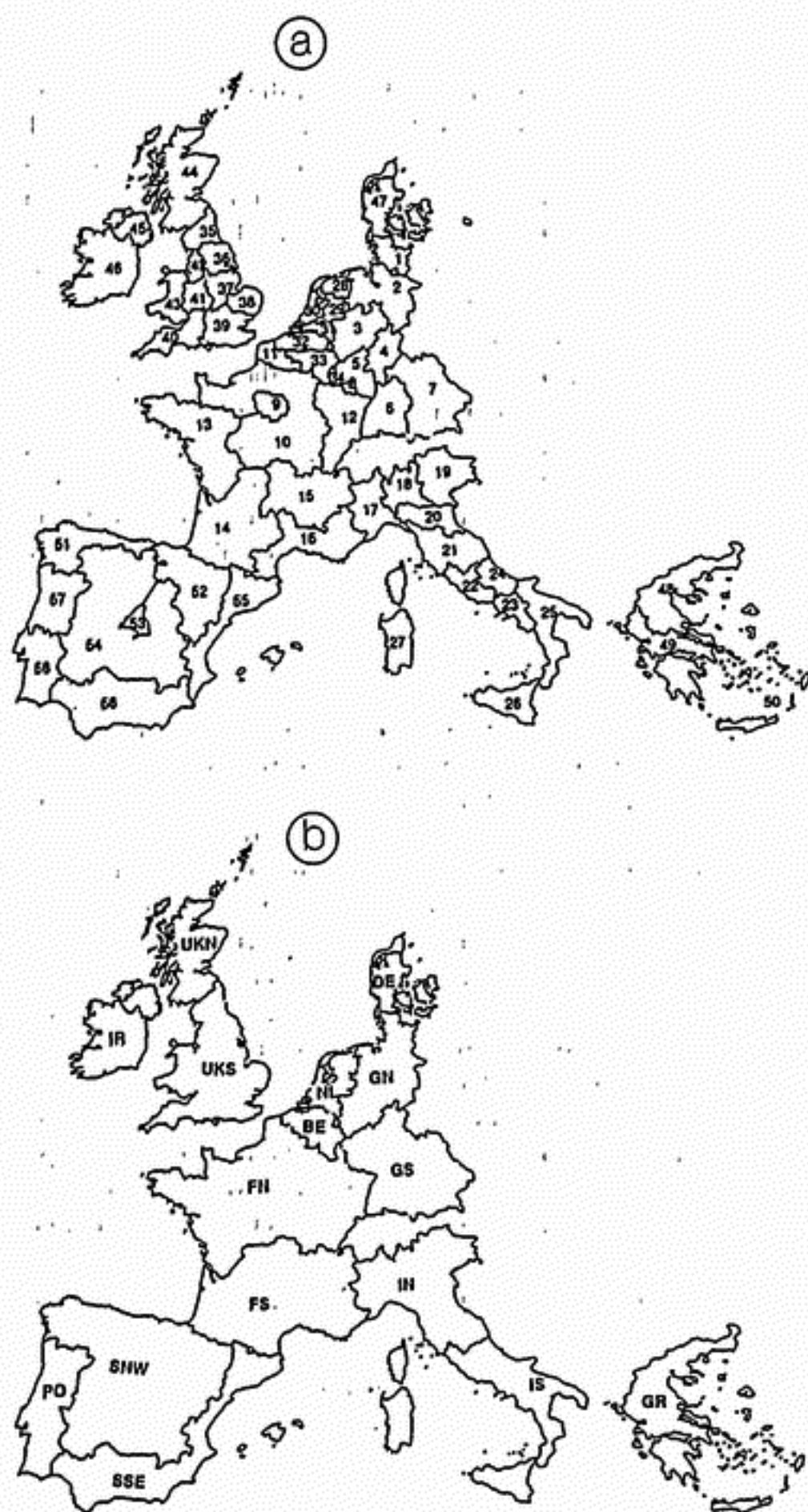


Figure 3. (a) – the NUTS-1 regions of the EC. (b) – the aggregated NUTS-1 regions ('NUTS-0' regions).

Spatial aggregation in GOAL

The model with land use activities and constraints defined at the NUTS-1 level was further aggregated to the member state (NUTS-0) level, with some exceptions: the five largest countries – UK, Germany, France, Italy and Spain – were subdivided into two parts (North and South), and Luxembourg was added to France North (Figure 3b). This resulted in a model with 16 spatial units ('NUTS-0' level).

The consequences of aggregating the NUTS-1 regions were analyzed in two steps. First, the regional (NUTS-1) resource constraints were aggregated (by addition) except for the area constraints. In physical terms this would imply the possibility of (free) transport of, for example, feed, manure and water across the borders of NUTS-1 regions. Area constraints were not aggregated in this step because the land use activities are NUTS-1 region-specific (soil and climate). After aggregation of the constraints, activities were aggregated as well, by calculating area weighted averages for the input-output coefficients. Table 3 presents the size of the various models. To analyze the effects of aggregating NUTS-1 regions, input-output combinations of land use activities at that level are considered first.

Aggregation of input-output combinations

Implications of aggregating NUTS-1 input-output combinations were investigated by plotting cumulative input against cumulative output with different regions ranked according to their output/input ratio. Figures 4a,b show the relation between crop production and total nitrogen loss and total water use for irrigation, respectively, in high-input irrigated silage maize production, before and after aggregation. The 58 NUTS-1 regions were aggregated to 16 regions in two ways: i) according to Figure 3b ('NUTS-0' level) and ii) in a random way. Input-output coefficients were aggregated by weighting the different regions according to their area suitable for maize production. Because the relation between nitrogen loss and yield is virtually linear, (i.e., the yield/nitrogen loss ratio is similar for all regions) both aggregation methods lead to a similar and small aggregation error. The relation between water use for irrigation and yield, however, is strongly curvilinear. After both aggregation methods, variation in the yield/water use ratio diminishes, the effect being much more pronounced after random aggregation than after 'NUTS-0' aggregation. In an optimization in which minimization of input use or input loss for a given total production is

Table 3. Number of activities and constraints in the NUTS-1 aggregated and 'NUTS-0' aggregated GOAL models.

Model	Activities	Constraints
NUTS-1 aggregated model	7834	1048
'NUTS-0' aggregation of constraints	7834	686
'NUTS-0' aggregation of constraints and activities	2206	323

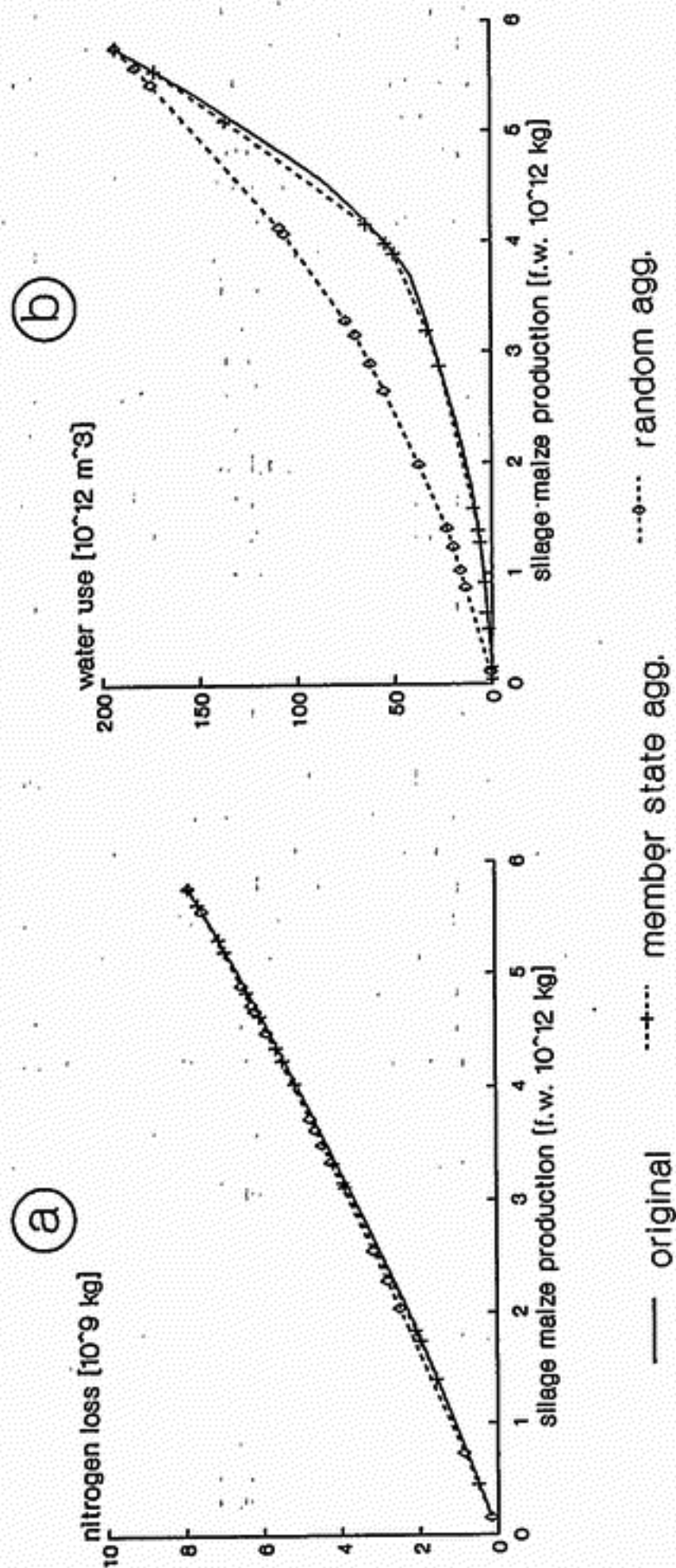


Figure 4. Cumulative nitrogen loss (a) or water use (b) against silage maize production in the EU before and after two ways of aggregation; regions have been ordered according to their output/input ratio.

the objective, the units with highest output/input ratios are selected. When variation in output/input ratio among different regions is high, random aggregation results in a large error in the objective values of an optimization, that is, a higher minimum input use for a given required production (Figure 4b). However, when activities with approximately the same output/input ratios are aggregated, this error remains small.

The relations in Figure 4 represent only two of many input-output relations in *GOAL* and are related to only some of the (conflicting) objective functions. For each combination of input and output, the relation (degree of curvilinearity and ranking of output/input ratios for the various regions) between cumulative input and cumulative output may be different. Moreover, it is not clear a priori which land use activities are selected in the optimal solution of the model. For instance, silage maize, which is used as animal feed, can be substituted by other roughage products or the area that could be used for efficient silage maize production may already be in use for the production of another crop. Aggregation of activities with linear input-output relations or aggregation of activities with similar output/input ratios yields a small aggregation error for the value of the objective function. However, because there are many land use activities in IMGLP models that are related to various objective functions, it is virtually impossible to predict the exact effect of aggregation by analyzing input-output relations as was done in Figure 4.

Effects of aggregation on optimum objective values at the EU level

Table 4 compares extreme solutions of the NUTS-1 and 'NUTS-0' aggregated models for five objective functions defined at the EU level. These extreme solutions are obtained by optimizing one objective function, without any upper or lower bound on the other objective functions. Differences in results are remarkably small. For all objectives, aggregation of constraints results in slightly better values (lower in minimizations, higher in maximizations; an increase of the feasible space) for objective functions. When aggregating activities and constraints, solutions become worse again. Compared with the NUTS-1 aggregated model, the 'NUTS-0' aggregated model shows only slightly worse extreme solutions. Hence, for extreme solutions of

Table 4. Minimum value of agricultural area [10^6 ha], total nitrogen loss [10^9 kg], total pesticide use [10^6 kg] or total costs of agriculture [10^9 ecu], or maximum value of agricultural employment [10^6 man-power units] in the NUTS-1 aggregated and 'NUTS-0' aggregated *GOAL* models, with no specific bounds on the objective functions that are not optimized.

Objective function	Model		
	NUTS-1 aggregated	Constraints 'NUTS-0' aggregated	Activities and constraints 'NUTS-0' aggregated
Agricultural area	31.0	30.5	31.3
Total nitrogen loss	2.21	2.15	2.23
Total pesticide use	31.5	31.2	31.6
Total cost of agriculture	98.1	97.7	98.9
Agricultural employment	3.14	3.18	2.50

objective functions, the 'NUTS-0' aggregated model gives a fair representation of the NUTS-1 aggregated model.

IMGLP land use models are usually applied to generate land use scenarios for a specific combination of objectives with upper or lower bounds. Both the NUTS-1 and 'NUTS-0' aggregated models were used to examine results of the following scenario: minimize costs of agricultural production, with upper limits on total nitrogen loss (2500 Gkg) and use of pesticides (35 Mkg) and no bounds on other objective functions. Effects of aggregation are, again, in agreement with the rules derived from the schematic example: aggregation of constraints only, results in a better value for the objective function that is optimized, whereas this value becomes worse again when the activities are aggregated as well (Table 5). The optimum objective values of the NUTS-1 and 'NUTS-0' aggregated models differ only slightly. These results are representative for all other scenarios that have been examined. The small aggregation errors for the optimum objective values can be explained by the fact that differences in output/input ratios between the various NUTS-1 regions within a 'NUTS-0' region are rather small.

Effects of aggregation on optimum land use allocation within the EU

Besides solutions for objective functions, the corresponding regional land use allocation is an important output of the IMGLP procedure. Figure 5a shows land use allocation at the NUTS-1 level of the NUTS-1 aggregated model for the scenario described in the previous paragraph. In Figure 5b these results of the NUTS-1 aggregated model have been aggregated to the 'NUTS-0' level, thus applying the principle 'first calculate, then average'. Figures 5c,d show the results of the 'NUTS-0' aggregated model for the same scenario. The differences between Figures 5c,d and Figure 5b demonstrate that the regional land use allocation is much more sensitive to the level of aggregation than the solution of the objective functions (Table 5). For instance, in Belgium hardly any land is used in the NUTS-1 aggregated model (Figure 5b), whereas 50–100% is used in the 'NUTS-0' aggregated model. In North Italy,

Table 5. Solution of five objective functions (total costs of agriculture [10^9 ecu]); total nitrogen loss [10^9 kg]; total pesticide use [10^6 kg]; agricultural area [10^6 ha]; agricultural employment [10^6 manpower units]; in the NUTS-1 aggregated and 'NUTS-0' aggregated GOAL models in a scenario (see text).

Objective function	Model		
	NUTS-1 aggregated	Constraints 'NUTS-0' aggr.	Activities and constraints 'NUTS-0' aggregated
Total cost of agriculture (minimum)	110.9	110.3	111.7
Total nitrogen loss (upper bound)	2500	2500	2500
Total pesticide use (upper bound)	35	35	35
Agricultural area (no bounds)	55.0	55.0	55.5
Agricultural employment (no bounds)	2.18	1.97	2.02

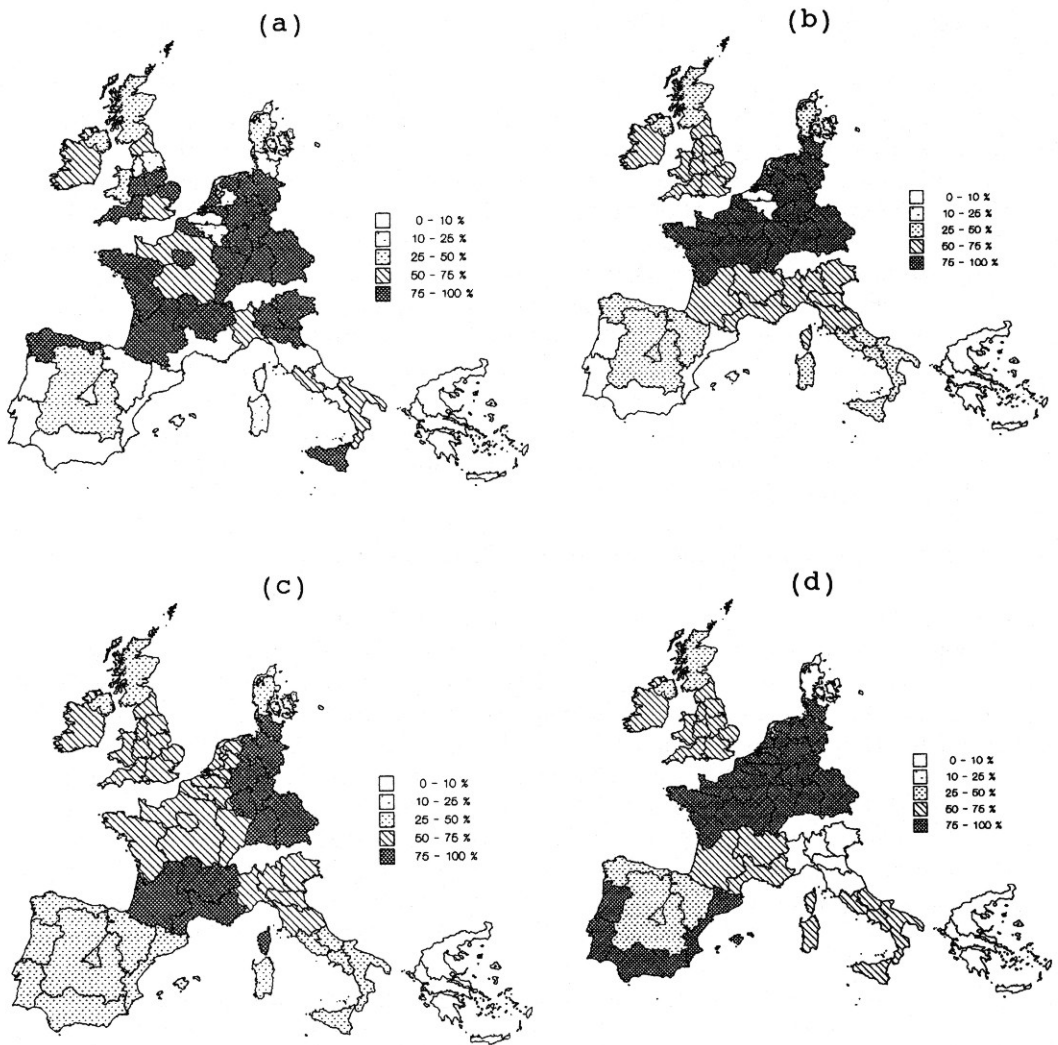


Figure 5. The area used for agricultural production (as a percentage of the suitable area) as a result of the *GOAL* model in the scenario presented in Table 5. (a) – results of the NUTS-1 aggregated model at the NUTS-1 level; (b) – ‘NUTS-0’ aggregated results of the NUTS-1 aggregated model; (c) – results of the model with ‘NUTS-0’ aggregated constraints; (d) – results of the model with ‘NUTS-0’ aggregated activities and constraints.

land use is considerably lower in the ‘NUTS-0’ aggregated model than in the NUTS-1 aggregated model.

In an minimization of an *IMGLP* model, land use activities with the highest output/input ratio (inputs or outputs referring to the objective function that is minimized) are selected. If different forms of land use in various regions have similar output/input ratios, values for the objective functions may differ only slightly under

very different land use patterns ('pin-table effect'). However, the ranking of the land use activities according to their output/input ratios may differ considerably between the NUTS-1 model and a more aggregated model. Hence, when optimizing an objective function, land use allocation may change considerably because of an aggregation, whereas optimum values of objective functions may change only slightly. In our example the output/input ratios of the 'NUTS-0' aggregated activities of the two NUTS-1 regions in Belgium (Vlaams-Gewest and Region-Wallone) become relatively more favourable than those of the 'NUTS-0' aggregated activities of the five NUTS-1 regions in North Italy (Nord-Ovest, Lombardia, Nord-Est, Emilia-Romagna and Centro). Consequently, agricultural production is allocated more to Belgium and less to North Italy.

Discussion and conclusions

In IMGLP models that explore land use options, results of quantitative land evaluations and analyses of agricultural production possibilities can be confronted with various social, economic and environmental objectives. We have analyzed the consequences of the spatial aggregation level in such a linear programming model. We discriminate between effects on objective functions which are defined generally at the level of the total system under study and effects on land use allocation within the system under study.

Effects on objective functions of aggregating spatial information were analyzed by discriminating between effects of aggregating activities and constraints in the LP model. Aggregation of activities reduces the feasible space for an LP model because spatial units with the highest (or lowest) output/input ratios are averaged with those having lower (or higher) ratios. Aggregation of constraints increases the feasible space – constraints become less binding because resources of two or more units can be then used in that region in which they are used most efficiently. Thus, aggregation of activities results in worse optimum values of objective functions, whereas aggregation of constraints results in better values. The schematized example and the examples with *GOAL* showed that the overall effect of aggregation on optimum objective values depends on 1) the degree of curvilinearity in input-output relations: the more curvilinear input-output relations are, the more serious the effects of aggregation can be; 2) the method of aggregation: aggregation of activities with similar output/input ratios has the smallest effects; 3) whether resource constraints are binding for the spatial units to be aggregated: aggregation has an effect only if constraints are binding; 4) the total area within the system that is used to meet the required level of agricultural production. Though these rules or factors can be used to analyze the effect of aggregating spatial information in LP models, it is practically impossible to predict their ultimate effect. This is particularly true in large LP models with multiple objectives, such as IMGLP models for land use studies. Thousands land use activities and constraints should be considered in relation to various objective functions.

In the aggregation from NUTS-1 to member-state ('NUTS-0') level in the *GOAL* model, activities with similar output-input ratios have usually been aggregated be-

cause differences in physical conditions within countries are generally smaller than those between countries. As a consequence, the aggregation error is rather small for the values of the objective functions at the EU level.

Besides the optimum values of the objective functions for the system as a whole, the corresponding land use allocation within the system is an important result in land use studies using the IMGLP procedure. The results obtained with the *GOAL* model show that the effect of aggregation on allocation of land use can be very important, even if the effects on the optimum values of the objective functions are small. If the output/input ratios of land use activities are similar for various spatial units (sub-regions), the values for objective functions are almost the same for rather different land use allocations. As has been shown in this paper, the effects of aggregation on land use allocation may be substantial because differences between output/input ratios of land use activities become smaller after aggregation and the ranking of spatial units according to their output/input ratios may change considerably.

In testing the sensitivity of the *GOAL* model for uncertainty in input-output coefficients, Scheele (1992) found that land use allocation is much more sensitive to this uncertainty than optimum values of objective functions. Scheele's results and our results strongly indicate that the results of land use models like *GOAL* are robust in terms of optimum values of objective functions defined at the level of the total system under study (in *GOAL*, the EU level) but much more fragile in terms of regional land use allocation within the system. The results at the regional level (in *GOAL*, e.g., the NUTS-1 level or member states) should be interpreted cautiously, and in relation to the level of spatial aggregation of information in the IMGLP model.

The complexity in predicting effects of aggregation on objective values and the important effects that aggregation may have on the regional land use allocation imply that aggregation of information in IMGLP models should be avoided. The motto 'first calculate, then average' (Rabbinge & Van Ittersum, 1994) also applies for land use studies using linear programming. In the *GOAL* model the NUTS-1 level is also an aggregated level, and although for instance East-Netherlands has not been selected in Figure 5a, it might be that when using less aggregated information in the LP model (e.g., at the NUTS-2 or NUTS-3 level) part of East-Netherlands (e.g., the Flevopolders) would have been selected in the optimum solution for the same or better objective values. The motto in exploratory land use studies is therefore 'first optimize the linear programming model and then aggregate to the appropriate policy level'. This motto is of course subject to practical limitations. Even with modern software and hardware, the size of LP models can become 'a binding constraint'. Information should then be aggregated according to agro-ecological criteria (land evaluation units with similar characteristics; i.e., similar output/input ratios for activities and similar resource constraints). After optimizing the LP model, the results can be translated into a spatial scale that is appropriate from a policy point of view.

Acknowledgements

We thank Drs D. Scheele for providing the original *GOAL* model and for the stimu-

lating discussions on simplification of this model. We acknowledge prof. H. Van Keulen, prof. R. Rabbinge and Dr F.R. Veeneklaas for their valuable comments on an earlier version of the manuscript.

References

- Day, R.H., 1963. On aggregating linear programming models of production. *Journal of Farm Economics* 45: 797-813.
- De Koning, G.H.J. & C.A. Van Diepen, 1992. Crop production potential of rural areas within the European Communities. IV: Potential, water-limited and actual crop production. Technical Working Document (W68), Netherlands Scientific Council for Government Policy, The Hague, 83 pp.
- De Koning, G.H.J., H. Van Keulen, R. Rabbinge & H. Janssen, 1995. Determination of input and output coefficients of cropping systems in the European Community. *Agricultural Systems* 48: 485-502.
- De Wit, C.T., 1993. Physiology and terminology. In: P.A. Leffelaar (Ed.), On systems analysis and simulation of ecological processes. Current Issues in Production Ecology, Volume 1, Kluwer Academic Publishers, Dordrecht, the Netherlands, pp. 3-9.
- De Wit, C.T. & H. Van Keulen, 1987. Modelling production of field crops and its requirements. *Geoderma* 40: 254-265.
- De Wit, C.T., H. Van Keulen, N.G. Seligman, & I. Spharim, 1988. Application of interactive multiple goal programming techniques for analysis and planning of regional agricultural development. *Agricultural Systems* 26: 211-230.
- Egbert, A.C. & H.M. Kim, 1975. Analysis of aggregation errors in linear programming planning models. *American Journal of Agricultural Economics* 57: 292-301.
- Nonhebel, S., 1994. The effects of use of average instead of daily weather data in crop growth simulation models. *Agricultural Systems* 44: 377-396.
- O'Neill, R.V. & B. Rust, 1979. Aggregation error in ecological models. *Ecological Modelling* 7: 91-105.
- Paris, Q. & G.C. Reuser, 1973. Sufficient conditions for aggregation of linear programming models. *American Journal of Agricultural Economics* 55: 659-666.
- Rabbinge, R. & M.K. Van Ittersum, 1994. Tension between aggregation levels. In: L.O. Fresco, L. Stroosnijder, J. Bouma & H. Van Keulen (Eds.), The future of the land: mobilizing and integrating knowledge for land use options. John Wiley & Sons Ltd., Chichester, pp. 31-40.
- Rabbinge, R. & H.C. Van Latesteijn, 1992. Long-term options for land use in the European Community. *Agricultural Systems* 40: 195-210.
- Rabbinge, R., C.A. Van Diepen, J. Dijsselbloem, G.J.H. De Koning, H.C. Van Latesteijn, E. Woltjer & J. Van Zijl, 1994. Ground for choices: a scenario study on perspectives for rural areas in the European Community. In: L.O. Fresco, L. Stroosnijder, J. Bouma & H. Van Keulen (Eds.), The future of the land: mobilizing and integrating knowledge for land use options. John Wiley & Sons Ltd., Chichester, pp. 95-121.
- Scheele, D., 1992. Formulation and characteristics of GOAL. Technical Working Document (W64), Netherlands Scientific Council for Government Policy, The Hague, 64 pp.
- Spreen, T.H. & T. Takayama, 1980. A theoretical note on aggregation of linear programming models of production. *American Journal of Agricultural Economics* 62: 146-151.
- Van Ittersum, M.K., R.J. Hijmans & D. Scheele, 1995. Description and user guide of GOAL-QUASI: an IMGLP model for the exploration of future land use. Quantitative Approaches in Systems Analysis, no 2. DLO Research Institute for Agrobiological and Soil Fertility and C.T. de Wit Graduate School for Production Ecology, Wageningen, 19 pp.
- Van Keulen, H., 1990. A multiple goal programming basis for analyzing agricultural research and development. In: R. Rabbinge, J. Goudriaan, H. Van Keulen, F.W.T. Penning de Vries & H.H. Van Laar (Eds.), Theoretical Production Ecology: reflections and prospects. Simulation Monographs, Pudoc, Wageningen, pp. 265-276.
- Van Lanen, H.A.J., C.A. Van Diepen, G.J. Reinds, G.H.J. De Koning, J.D. Bulens & A.K. Bregt, 1992. Physical land evaluation methods and GIS to explore the crop growth potential and its effects within the European Communities. *Agricultural Systems* 39: 307-328.

- Vecencklaas, F.R., S. Cissé, P.A. Gosseye, N. Van Duivenbooden & H. Van Keulen, 1991. Competing for limited resources: the case of the fifth region of Mali. Report 4: Development scenarios. CABO/ESPR, CABO, Wageningen, 180 pp.
- Zipkin, P.H., 1980a. Bounds on the effect of aggregating variables in linear programs. *Operations Research* 28: 403-418.
- Zipkin, P.H., 1980b. Bounds for row-aggregation in linear programming. *Operations Research* 28: 903-916.