IIASA
International Institute for Applied Systems Analysis • A-2361 Laxenburg • Austria
Tel: +43 2236 807 • Fax: +43 2236 71313 • E-mail: info@iiasa.ac.at • Web: www.iiasa.ac.at

INTERIM REPORT IR-98-026/May

Modelling producer decisions on land use in a spatial continuum.

Michiel Keyzer (michiel@sow.econ.vu.nl)
Yuri Ermoliev (ermoliev@iiasa.ac.at).

Approved by
Günther Fischer (fisher@iiasa.ac.at)
Leader, *Land Use Change Project*

Contents

Abstract

The paper describes how stochastic optimization techniques can be used to model profit maximizing producer behaviour in a spatial continuum. The main methodological issues to be addressed are, first, that the representation of optimal allocations in a spatial continuum naturally leads to models that contain integrals over space, and the second that the resulting model tends to have a multi-level structure, i.e. requires solving nested optimization problems because it should combine the profit maximization by individual producers with market clearing at regional level. We specify four regional models that may illustrate the approach. The first determines the optimal output level for factories that emit pollutants which reduce the crop output of neighboring farmers. The main issue is to compute the associated level of compensation to be paid by the factories to the farmers. The second model deals with optimal zoning. It computes the optimal crop routing for farmers who can choose to sell their crop to factories situated at given locations. This is an optimization problem in functional space, which can be reformulated as a dual stochastic optimization problem. In the third model, the farmer has the possibility of routing his crop along different roads or distribution nodes to the various factories for processing. It can describe the optimal choice of distribution centres at given locations, around plants or cities, and produces optimal boundaries for the zones that supply to or buy from these centres. The fourth model deals with the problem optimal land consolidation, distinguishes between consolidation processes with and without side-payments. To each of these four models we associate a decentralised, stochastic quasi-gradient (SQG-)procedure for attaining the (global) optimum, which has a natural interpretation as a device for decentralized adaptive planning.

About the Authors

Professor Michiel Keyzer is Director of the Centre for World Food Studies, Vrije Universiteit (SOW-VU), Amsterdam, The Netherlands.

Dr. Yuri Ermoliev is a senior Research Scholar at IIASA.

Modelling producer decisions on land use in a spatial continuum

Michiel Keyzer and Yuri Ermoliev[1]

1. Introduction

The growing concern about environmental and social sustainability of current land-utilization and land-cover patterns has led to a revival of interest in problems of optimal spatial allocation. So far, the wealth of geographical data on climate, soils, land cover and groundwater flows that increasingly become available through remote sensing, has mainly been incorporated in Geographical Information Systems, which can produce highly detailed maps but so far contain only few and very simple decision support tools to optimize over a spatial continuum. As a rule, the geographically explicit decision models either describe a separate optimal decision at every point on the map without accounting for spatial interdependencies, or limit themselves to optimization over a relatively small number of regional units, severely at the expense of geographical detail.

The optimization over a spatial continuum has been studied in location theory, in connection with the problem of optimal facility location. This problem usually amounts to finding the geographical location of an industrial facility that minimizes the cost of transporting goods to that location from a surrounding region or vice versa. The early location models are classical transportation models and only select the best out of a finite number of alternatives and treat the region as a finite number of fields identified by their barycentre (see e.g. Beck and Goodin, 1982, for an application to dairy farms). Subsequent location/pricing models treat the site as a continuous choice variable and simultaneously calculate the consumer price at every point in the region on the basis of the distance from the facility (Hansen et al., 1987, and Drezner, 1995, for a survey).

However, the question of optimal land use requires a broader treatment. It calls for an explicit spatial representation of land use itself, with land being allocated to competing uses so as to maximize, say, the total revenue in the region, as opposed to the cost minimization of the location/pricing models. Consequently, from a modeling point of view new methodological challenges must be addressed. The first issue is that the representation of optimal allocations in a spatial continuum naturally leads to models that contain integrals over space, which in general cannot be eliminated analytically, in view of the variability in the underlying GIS-information. The second issue is that the resulting model, which combines optimal routing and profit maximizing decisions by individual producers, with market clearing at regional level, tends to have a multi-level structure, i.e. requires solving nested optimization problems.

---

[1]The first author is affiliated with the Centre for World Food Studies, Vrije Universiteit Amsterdam, the second author with the International Institute for Applied Systems Analysis, Laxenburg, Austria, and the Academy of Sciences of the Ukraine.

The aim of this paper is to show that stochastic optimization techniques allow to address both issues. We specify four regional models in which profit maximizing producers are operating within a spatial continuum and compete on land and commodity markets. To every model we associate a decentralised, stochastic quasi-gradient (SQG-)procedure for attaining the (global) optimum (see Ermoliev, 1988). Since in this paper we are more interested in such algorithms as practical devices for decentralized adaptive planning than as efficient tools for computations, we focus on interpretation and only refer in passing to operational issues such as the efficient choice of step-size, or the speed of convergence.

The paper proceeds as follows. In section 2 we describe by means of simple examples how stochastic optimization models can be used for planning problems in the spatial continuum. Next, we deal with applications of increasing complexity. First, in section 3, we study a model that determines the optimal output level for factories that emit pollutants which spread over the neighbouring environment, and reduce the crop output of local farmers. The main issue is to compute the associated level of compensation to be paid by the factories to the farmers. The technical difficulty is to deal with an integral in the objective of a convex program. Our second problem (section 4) deals with optimal zoning. It computes the optimal crop routing for farmers who can choose to sell their crop to factories situated at given locations. We indicate that this zoning aspect generates an optimization problem in functional space, which can be reformulated as a dual stochastic optimization problem. The model of section 5 enables the farmer to route his crop along different roads or distribution nodes to the various factories. This model can describe the optimal choice of distribution centres at given locations, around plants or cities, and produces optimal boundaries for the zones that supply to or buy from these centres. Next, we reinterpret this model as a planning tool for optimal land consolidation and describe how the solution path generated by the SQG-algorithm can interpreted as a sequence of land transactions with side-payments, that eventually converges to the optimum. However, in many practical situations it will be unrealistic to suppose that such side payments can be mobilized, because individual farmers are often reluctant to participate in land transactions if their farm becomes less profitable in the process. To deal with this case, we formulate an alternative model that maximizes the revenue without side payments of the least favoured. Section 6 concludes.

2. Solving spatial planning problems by stochastic optimization

To illustrate the difficulties that arise when modeling the decisions in a spatial continuum, let us consider some typical examples. The calculation of the barycentre of given geographic region is a well known case. One might think of the possibly dispersed farmland whose "centre" has to be determined, say, in order to serve as the collection point for the harvest. Let X denote the region and $A \subseteq X$ the (not necessarily connected) land of a given farmer. The barycentre of A can now be calculated by minimizing

$$F(h) = \int_A \|x-h\|^2 \, dx \tag{2.1}$$

where $\|\cdot\|$ is the Euclidean distance, which is equivalent to the stochastic optimization problem of minimizing the expectation function $F(h) = \int_A \|x-h\|^2 \, dG(x)$, where $G(x) = x / \int_A dx$ is a probability measure. This is an important point, as it connects spatial planning problem to stochastic optimization.

The solution[2] $h^* = \int_A x\, dx / \int_A dx$ can be computed by discretization, numerical integration, either directly, or via sequential Monte Carlo simulation. Discretization is common in GIS-packages. It approximates (2.1) by a grid of N points x^j, and computes the barycentre as the average coordinate, $h = \Sigma_j x^j/N$. Sequential Monte Carlo simulation proceeds as follows. Sample at iteration t = 1, 2, ... a point x(t) at random and independently, such that $\text{Prob.}[x^t \in A] = \int_A dx / \int_A dx$, and define the approximate solutions

$$h(t+1) = h(t) - \rho_t\,(h(t) - x(t)), \qquad t = 1, 2, ... \tag{2.2}$$

Notice that (2.2) is actually an SQG-algorithm, since $2(h(t)-x(t))$ is the gradient of the random (sample) function under the sign of the integral in (2.1). If the stepsize ρ_t satisfies:

$$\rho_t \geq 0,\ \Sigma_{t=1}^{\infty} \rho_t = \infty,\ \Sigma_t^{\infty}{=}1\ \rho_t^2 < \infty, \tag{2.3}$$

then the sequence {h(t)} will converge to h*. This requirement will be met for $\rho_t = 1/t$, which corresponds to the sample mean calculation:

$$h(t) = 1/t\ \Sigma_{k=1}^{t}\ x(k). \tag{2.4}$$

The advantage of (2.2) as opposed to (2.4) is that it does not require to store a large bundle of information, since in (2.2) the estimation of h* proceeds sequentially, with a simple updating rule after each new observation. Procedure (2.2) is actually an SQG-algorithm for minimizing the integral (2.1). However, the squared Euclidean distance will in general not be the appropriate criterion for the profit maximizing producer, who is interested in, say, minimization of average distance itself rather than average squared distance. Moreover, it allows for infinitely many routes from point x to the variable "home" h*. It seems more realistic to treat this home or market outlet as fixed (let this be point b for base) and the farmer to choose the shortest (or cheapest) route, by searching for collection points h^i, i = 1, ..., r.

The associated zoning pattern defined by a function i(x) which indicates that point i attracts x: $i(x) = \text{argmin}_i\ \|x-h^i\| + \|h^i-b\|$. This function can be shown in a GIS zone-map. Other "statistics" could include, for example, the total cost

$$z_0 = \int_A \min_i\ [\|x-h^i\| + \|h^i-b\|]\ dG(x), \tag{2.5}$$

where G(x) might reflect the demand for transport i.e. the crop output at point x) and the traffic on the various roads:

$$z_i = \int_A \delta_i(x)\ dG(x),\ i = 1, ..., r$$

---

[2] Calculation is coordinate-wise: $h_1 = \int_A (x_1, 0)\, dx \,/\, \int_A (1, 0)\, dx$ (1.2)
$h_2 = \int_A (0, x_2)\, dx \,/\, \int_A (0, 1)\, dx$.

where $\delta_i(x) = 1$ if $i(x) = i$ and 0 otherwise. These values $z_0, z_1, ..., z_r$ can be estimated by discretization but also as minimizers of

$$F(z) = (z_0 - \int_A \min_i [\|x-h^i\| + \|h^i-b\|] \, dG(x))^2 + \sum_i (z_i - \int_A \delta_i(x) \, dG(x))^2. \qquad (2.6)$$

The associated SQG-procedure, which is similar to (2.2), will sequentially process a large amount of information and with probability 1 yield a consistent estimate of z. The examples discussed can all be dealt with relatively easily by alternative discretization methods, because their first-order optimality conditions yield an explicit solution for the unknown parameters q. We now turn to models that do not possess this special property.

3. Optimal land use in the presence of pollution

We consider industrial plants whose pollution negatively affects neighbouring land users. If the emissions consisted only of pollutants such as CO_2-gases that tend to dissipate quickly into the atmosphere, the analysis could focus on reduction of aggregate emissions, and there would be no need for a locational study. However, most emissions have definite local component whose cumulative effects depend on location specific factors such as soil type, hill slope, crop coverage and climatological factors.

The spreading of cumulative emissions of a given pollutant is naturally represented via two-dimensional density functions. These are defined over the geographical region under consideration and measure the incidence (fraction) of pollutant emitted by a factory located at given site that depositions at every point in the region. The factory might be a chemical plant or an intensive livestock farm. The decision problem is to confront every factory with the land users in its neighbourhood (actually in its environment) who suffer from the emissions, i.e. to maximize total income or welfare of the region while internalizing the environmental effects, and this naturally leads optimization over a spatial continuum. Polluters will pay compensation to the land users. This will affect their profitability, and size of operation at every site, as well as the revenues and cropping patterns of the land users around them.

Model formulation

The model is specified as follows. We denote the given geographical space of region by the set $X \subset R_+^2$ and consider S fixed factory sites indexed s, located at $b^s \in X$. A factory at site s makes use of an emission permit for a quantity q_k^s of pollutant k, for k = 1, ..., K. This use of permits is taken to increase its revenue $R^s(q^s)$ from the factory like an ordinary input; formally:

Assumption R (Revenue function of factory at fixed site s): For every site s, the revenue function $R{:}R_+^K \to R$, $R^s(q^s)$ is increasing, continuously differentiable with respect to q^s, with uniformly bounded derivative, and strictly concave in the input q^s.

The emissions are dispersed around the factory by various physical processes such as winds and groundwater flows. Let the density function $\psi_k^s(x)$, defined over X describe the dispersal of emissions according to $c_k(x) = \Sigma_s q_k^s \psi_k^s(x)$, the incidence of pollutant k at location x, where x is a two- or three-dimensional vector. This reduces environmental quality at point x, leaving less natural inputs g(x) = ω(x) - c(x) for crops, where ω is the given resource availability, and the revenue from crop farming at location x will be affected. The local revenue function r(g, x), for which satisfies:

Assumption r (revenue function of farm at point x in region X): the revenue function $r: R^K \times X \to R$, r(g, x) is integrable in x and, almost everywhere on X, continuously differentiable w.r.t. the input g, with uniformly bounded derivative, and strictly concave in g.

In assumptions R and r, the requirement of continuous differentiability can be relaxed, but it will be seen to offer the advantage of ensuring uniqueness of the market clearing prices. The concavity implies that there is no possibility of polluting without further damage an already fully spoiled area. Allowing for increasing returns in pollution would undermine concavity. Yet the formulation allows to consider wasteland whose revenue does not dependent on pollution. Concavity is taken to be strict to keep the optimal q* unique. Finally, in assumption r, integrability in x is obviously necessary for integral calculations, while uniform boundedness of the derivative to q is a requirement of the SQG-procedure itself, which could be relaxed but is maintained for simplicity. Below we will comment on the qualification "almost everywhere".

The problem is now to define the optimal production levels, while maximizing the revenue in the region and to determine the compensatory payments of the factories to the other users of environmental resources. The revenue maximizing regional model is:

$$\max_{q^s \geq 0,\ \text{all } s,\ g(x) \geq 0,\ \text{all } x} \ \Sigma_s R^s(q^s) + \int r(g(x), x)\, dG(x)$$

subject to (3.1)

$$g_k(x) + \Sigma_s q_k^s \psi_k^s(x) \leq \omega(x), \qquad \text{for } k = 1, ..., K, \text{ all } x,$$

where G(x) is a distribution on the domain X, and the symbols under the maximization denote choice variables. This is a convex program, since the objective is concave and the constraint set linear. It has an infinite number of constraints. Let us briefly characterize the solution of this problem. We assume that g(q, x) is everywhere on X positive in the optimum. Since $R^s(q^s)$ is increasing in q^s and the density is nonnegative, the constraint holds with an equality and we can define $g_k(q, x) = \omega(x) - \Sigma_s q_k^s \psi_k^s(x)$, and the resulting first-order conditions are:

$$\partial R^s(q^s)/\partial q_k^s = -\int (\partial r(g(q, x), x)/\partial g_k)\ \psi_k^s(x)\, dG(x), \text{ if } q_k^s > 0. \qquad (3.2)$$

Proposition 1 (*Compensation payments by polluters*): Under assumption R and r, the regional revenue maximization (3.1) implies that

(1) every factory s pays a total transfer $\Sigma_k\ \partial R^s(q^s)/\partial q_k^s\ q_k^s$ to farmers;

(2) every spot x receives $\Sigma_k \ \partial r(g(q, x), x)/\partial g_k \ \psi_k^s(x) \ q_k^s$.

Proof. This follows directly from first-order optimality conditions.□

From this proposition follows that the receipts of a crop farmer from factory s will increase with (i) the emission q_k^s, (ii) the dispersion density $\psi_k^s(x)$ to location x, and (iii) the marginal damage $\partial r(g, x)/\partial g_k$ at location x. Conversely, factory s will have to restrict its emissions (and possibly even close down) if its pollution dissipates to locations x where the damage is important, either because the location itself is vulnerable or because other factories can pollute it less harmfully (i.e. can obtain more revenue from a marginal unit of pollution q_k^s. Therefore, the model can be interpreted as a location model, even though location b^s is fixed.

Solution procedure

The model can be solved by an SQG-process. Define the function

$$f(q, x) = \Sigma_s \ R^s(q^s) + r(\omega(x) - \Sigma_s \ q_k^s \psi_k^s(x) g(x), x), \tag{3.3}$$

and gradient

$$\partial f(q, x)/\partial q_k^s = \partial R^s(q^s)/\partial q_k^s + \partial r(g, x)/\partial g_k) \ \psi_k^s(x). \tag{3.4}$$

The spatial optimization problem (3.1) is now to maximize

$$F(q) = \int f(q, x) \ dG(x), \tag{3.5}$$

on the compact convex set Q, which is taken to be specified as $Q = \{q \mid 0 \le q \le q\}$. The SQG-algorithm considers a sequence of random drawings x(t) from the set X and starting from a given q(1) $= q^1 \in Q$, adjusts q(t) according to:

$$q(t+1) = \Pi_Q(q(t) + \rho_t \ \partial f(q(t), x(t))/\partial q), \quad t = 1, 2, \ldots \tag{3.6}$$

where Π_Q is the projection operator on Q (i.e. the point in Q nearest to the projected point), and the scalar step-size ρ_t converges to zero according to some appropriate step-size rule. From the general results on the convergence of SQG-methods follows that if f(q, x) is differentiable and concave in q, almost everywhere on G(x), then process (3.6) converges, with probability 1, to the optimal solution of (3.5), for stepsize ρ_t chosen as in (2.3).

As an alternative to the SQG-method one might consider using deterministic techniques after having approximated the function F(q) by its sample mean:

$$F^N(q) = 1/N \ \sum_{k=1}^N f(q, x(k)) \tag{3.7}$$

where x(k) are independent samples from G(x). However, for most of the models discussed in this paper the value N would have to be extremely large, and this would require a large number of terms f(q, x(k)), which might, moreover, be highly non-smooth and not available explicitly. For the SQG-procedure (3.6), the convergence principle can be understood as follows. Since the direction $\partial f(q^t, x^t)/\partial q$ coincides "on average" with the gradient $\partial F(q)/\partial q$, i.e.

$$\int \partial f(q(t), x(t))/\partial q)\ dx = \partial \int [f(q(t), x)\ dx]/\partial q, \tag{3.8}$$

on average the value F(q) increases from one iteration to the next, and for a step-size moving to zero at appropriate speed, the sequence {q(t)} generated by adjustment rule (2.2) will converge to the optimal solution. In fact this rule can be viewed as a stochastic decentralisation procedure. It has an "evolutionary" interpretation as a decentralised learning process. Once a point x^t has been selected by at random within the region (the mutation), the vector $q(t) \in Q$ gradually changes its composition (the selection) so as to improve performance (survival). Whereas a deterministic gradient method would use the gradient $\int \partial f(q(t), x)/\partial q)\ dG(x)$ and thus assume the distribution G(x) to be known to the planner, this SQG-adjustment only improves fitness from the perspective of the randomly chosen point x(t), without any use of information on the distribution or on the value F at any other point x. Thus, the process is perfectly decentralised and derives its coordination (optimality of allocations) from the infinite repetition of events and a relatively "naive" adjustment to current pressure as expressed through the quasi-gradient. And, while the single point has a negligible influence on the final outcome, the optimal solution is eventually determined by the shape of the distribution G(x), jointly with the function f(q, x) and the set Q.

Let us now return to the qualification, "almost everywhere" in conjunction with the "integrability in x" of Assumption r. Both are significant relaxations of the usual demands that differentiability and concavity properties should apply everywhere on the domain. This is important in a spatial context, as they allows for the representation of natural discontinuities due to, say, rivers or roads. Technically, the role of this qualification is expressed in the following simple lemma:

Lemma 1 (Mean value) Consider the distribution G(x) and the function f: Q×X, f(q, x), which is integrable w.r.t. x. If, almost everywhere (the points where this does not hold have measure zero), f(q, x) is concave w.r.t. q then $F(q) = \int f(q, x)\ dG(x)$ is concave on Q.

Proof. Choose two arbitrary points q_1 and q_2 from Q and consider an arbitrary convex combination $q_3 = \lambda q_1\ (1-\lambda)q_2$, for any $\lambda \in [0,1]$. Then, by concavity w.r.t. q, $\lambda\ v(q_1, x) + (1-\lambda)\ v(q_2, x) \le v(q_3, x)$ almost everywhere on G(x). Therefore, $\lambda \int f(q_1, x)dG(x) + (1-\lambda) \int f(q_2, x)dG(x) \le \int f(q_3, x)dG(x)$ and F(q) is concave on Q.□

The representation in a continuum has special attraction for spatial problems because, once an optimal value q* has been estimated, GIS-tools allow to produce, say, an "altitude" map of the farm revenues $r(\omega(x)-\Sigma_s\ q_k^s{*}\psi_k^s(x), x)$. Furthermore, once problem (3.5) has been solved, the results can be compared in a spatially explicit manner with those without compensation where the permits have zero price and the factories maximize $R^s(q^s)$, while the land-users are confronted with given pollution levels.

4. Zoning problem

The problem of assigning land to factories located at given sites is known as zoning. Instead of describing a physical flow of pollutants from a point s to a surface, the zoning problem assigns surfaces to given locations. Assume that farm output is carried from the fields to the collection point, which might be a factory, but also a city and even the farmer's homestead. Every field is assumed to grow the most profitable crop (combination) at every spot and the harvest is carried to a central site to be chosen so as to yield the highest revenue after accounting for transportation costs. Thus, fields are optimally associated to sites.

Model formulation

The variable q of model (3.1) will, as before, be the input into the factory and will now represent a physical flow $y_k^s(x)$ of harvested crop k from point x to site s, located, say, at point b^s. The model will, as in (3.1) have an integral in its objective but, in addition, the objective will contain the spot-specific endogenous variable $y_k^s(x)$. This introduces the difficulty that the exogenously given density function should be replaced by an endogenous routing decision. Transportation costs are denoted by $w_k^s(x)$ a function that should be integrable on G(x).

Transportation costs might be a function of the Euclidean distance between x and b^s. The formulation could easily be generalized to account for possible asymmetry in transportation costs (up-hill/down-hill), but here we disregard this aspect for convenience. Let q^s now denote the aggregate flow of crop output to point b^s, to be used as input by the factory, and let the revenue function $R^s(q^s)$ satisfy Assumption R. We write $r(-\Sigma_s y^s(x), x)$ for the associated land revenue function (that transforms crop output into revenue), which satisfies Assumption r. Finally, let p_k^s denote the purchasing price that clears the commodity market k at site s, whose balance reads:

$$q_k^s - \int y_k^s(x)\, dG(x) \leq 0, \text{ all k and s.} \tag{4.1}$$

To represent this model formally, we proceed in the way that is usual in location analysis and replace the unknown supply function $y_k^s(x)$ by a well specified function of prices, as follows. For a given value p^s define the profit functions:

$$\Pi^s(p^s) = \max_{q^s \geq 0} (R^s(q^s) - \Sigma_k\, p_k^s q_k^s), \qquad \text{all s} \tag{4.2a}$$

which is convex non-increasing in input price p^s. For spot x, the profit from routing the crop to site s, is:

$$\pi^s(p^s - w^s(x), x) = \max_{y^s(x) \geq 0} \Sigma_k\, (p_k^s - w_k^s(x)) y_k^s(x) + r(-y^s, x), \tag{4.2b}$$

which is convex nondecreasing in output price $(p^s - w^s(x))$. Competitive market prices p^s solve the master problem:[3]

$$\min_{p^s \geq 0,\ \text{all } s} \Sigma_s \Pi^s(p^s) + \int \max_s \pi^s(p^s - w^s(x), x)\ dG(x). \tag{4.3}$$

Continuous differentiability in Assumption r implies strict convexity in p of this objective (see Ginsburgh and Keyzer (1997), ch. 2), while strict concavity in assumption R ensures continuous differentiability of $\Pi^s(p^s)$. Hence, by Hotelling's lemma, optimal inputs q^s satisfy

$$q^s = -\partial\Pi^s(p^s)/\partial p^s. \tag{4.4}$$

while, almost everywhere on X, $\pi^s(p^s - w^s(x), x)$ is continuously differentiable, and outputs y^s are given by:

$$y^s(x) = \delta^s(x)\ \partial\pi^s(p^s - w^s(x), x)/\partial p^s, \tag{4.5}$$

where $\delta^s(x) = 1$ if routing to s yields maximal profit, and 0 otherwise. Notice that by construction, the routing will be the same for all commodities produced at point x. Finally, the first-order optimality conditions of (4.3) ensure satisfaction of the commodity balances:

$$\int y^s(x)\ dG(x) = q^s. \tag{4.6}$$

We notice also that by the conjugate function theorem, the problem in quantity terms dual to (4.3) is:

$$\begin{aligned} &\max_{q^s \geq 0,\ y^s_k(x) \geq 0 \text{ all } k,\ s, \text{ all } x} \Sigma_s R^s(q^s) + \int (r(-\Sigma_s y^s(x), x) - \Sigma_s \Sigma_k w^s_k(x) y^s_k(x))\ dG(x) \\ &\text{subject to} \\ &q^s_k - \int y^s_k(x)\ dG(x) \leq 0, \text{ all } k \text{ and } s \end{aligned} \tag{4.7}$$

which is a problem in functional space.

Solution procedure

Let us now turn to the solution procedure for model (4.3). The problem is to minimize

$$V(p) = \int v(p, x)\ dG(x) \tag{4.8}$$

on a compact convex price set $\mathcal{P} = \{0 < \underline{p} \leq p \leq \overline{p}\}$, for $v(p, x) = \Sigma_s \Pi^s(p^s) + \max_s \pi^s(p^s - w^s(x), x)$. Due to the $\max_s$-operator, the function v(p, x) is non-smooth but under assumption r, it is almost

[3] Although model (4.3) only deals with farm output, it is relatively straightforward to account for inputs, by extending the profit function to $\Pi^s(p^s, p^s)$ and $\pi(p^s - v^s(x - x^s), p^s + v^s(x - x^s), x)$ and requiring it to be, almost everywhere on X, jointly convex in the first two arguments, increasing in output and decreasing in inputs, and continuously differentiable in x.

everywhere on X, convex in p_s and any stationary point is a global optimum.

The solution procedure can be cast into an SQG-form for (decentralized) price adjustments. At step t = 1, 2, ..., the change of prices from p(t) to p(t+1) is initiated by the random choice of a location x(t) from the distribution G(x). This changes decreases the aggregate profit function $v(p, x) = \Sigma_s \Pi^s(p^s) + \max_s \pi^s(p^s - w^s(x), x)$ at x(t):

$$p(t+1) = \Pi_{\mathcal{P}}(p(t) - \rho_t\, \partial v(p(t), x(t))/\partial p), \qquad t = 1, 2, ... \tag{4.9}$$

where $\Pi_{\mathcal{P}}$ is the projection operator on $\mathcal{P}$. In fact, procedure (4.9) defines a stochastic Walrasian tatonnement of the type described in Ermoliev et al. (1997). Recall from (4.4) and (4.5) that since net supply is the derivative of the profit function $-\partial v(p(t), x(t))/\partial p$ is simply the net demand of point x(t), where the net demand from factories s is taken to be spread uniformly over the surface.

It is straightforward to introduce a spatial spread, say, via the density function $g^s(x)$, for the activities of the factory also. The criterion then becomes $v(p, x) = \Sigma_s \Pi^s(p^s)\, g^s(x) + \max_s \pi^s(p^s - w^s(x), x)$. This modification will affect the SQG-path but not the eventual optimum, because this density is fixed and does not affect commodity balance (4.1). As in the deterministic Walrasian tatonnement, the procedure changes prices in the direction of net demand (but here this if not the aggregate net demand), until aggregate (i.e. expected) net demand becomes zero eventually and markets are cleared.

Zoning

The solution of model (4.5) will yield market clearing prices p^s, from which input demands q^s by factories and farm supplies $y_k^s(x)$ can be recovered via (4.4) and (4.5), respectively. This raises the question as to the properties of the optimal values $y_k^s(x)$. Indeed, under relatively mild assumptions about transportation costs $w_k^s(x)$, and the differentiability of revenue function r(g, x), it is possible to show that almost everywhere on X, $y_k^s(x)$ will be i.e. be nonzero for at most one s and that it will be the same for all outputs k. This amounts to asserting that the model generates a specialized zoning, i.e. an integer valued mapping s(x) that is single valued, almost everywhere on x, thus creating a unique association between spots and factories.

Assumption W (unit cost of transportation from x site s): The cost of transporting one unit of commodity k to site s, $w_k^s(x)$ is, almost everywhere on X, (i) nonnegative and continuously differentiable in x. Moreover, (ii) almost everywhere on X and for every k and pairs [s,s′], s≠s′, the difference $c(x) = w_k^s(x) - w_k^{s'}(x)$ has non-zero gradient whenever c(x) = 0.

Assumption r2 (differentiability) the revenue function r(g, x) is almost everywhere on X continuously differentiable in x.

Various specific formulations of the cost functions, such as the Euclidean distance, guarantee satisfaction of assumption W(ii). Geometrically, this is because two circles with different centres can only share points but no arcs. We can now state and prove:

Proposition 2 (Specialization of spots): Let assumptions R, r and W hold. Then, (a) program (4.3) partitions the region X into zones with a common routing such that (a) no zone will produce for more than one factory, and (b) the number of zones is finite.

proof. Continuous differentiability of the revenue functions ensures that program (4.3) has a unique solution p*. Now for an arbitrary pair of sites [s,s′], s≠s′, define the function $z(x) = \pi^s(p^{s*} - w^s(x), x) - \pi^{s'}(p^{s'*} - w^{s'}(x), x)$. By assumptions W(i) and r2, the function z(x) is continuously differentiable almost everywhere on X. Also partition x into (x_1, x_2), where x_1 refers to the first coordinate of x (a scalar) and x_2 the other coordinate (we assume for convenience that the space is two-dimensional). The proof proceeds in two stages.

First, we keep x_2 fixed at an arbitrary value x_2^0 in X and characterize the fixed points for x_1, as defined by $z(x_1, x_2^0) = 0$, as follows. One possibility is that there are no solutions. This means that either $z(x_1, x_2^0) > 0$ (meaning that s dominates s′) or $z(x_1, x_2^0) < 0$ (s′ dominates s) for all points on the line $x_2 = x_2^0$. If solutions exist, these can be of two kinds: isolated or on closed segments along the line $x_2 = x_2^0$. Since X is compact the number of closed segments must be finite, and the fixed point counting theorem in Ortega and Rheinboldt (1970, p. 150) says that, because z(x) is continuously differentiable and X is compact, the number of isolated solutions is finite as well. We can therefore construct around the line $x_2 = x_2^0$, a band consisting of neighbourhood sets $N_j(x_2^0)$, indexed j = 1, ..., J that fully cover this line and do not overlap.

Secondly, consider the perturbed fixed points close to x_2^0, say, at $x_2^0+\varepsilon$ within the band. If there where no fixed points at x_2^0 this may still be the case at $x_2^0+\varepsilon$ but this only means that the domination persists. Otherwise $z(x_1, x_2^0+\varepsilon) = 0$ might have some fixed points. Consider in the neighbourhood $N_j(x_2^0)$ of the j-th solution at x_2^0, the set of perturbed solutions $Z_j(x_2^0) = \{x_1 \mid z(x_1, x_2^0+y) = 0, 0 \le y \le \varepsilon, (x_1, x_2) \in N_j(x_2^0)\}$. If this segment persisted, the set $Z_j(x_2^0)$ would have positive surface but would contradict assumption W(ii). If it does not persist, it can become an intersection or disappear altogether. In both cases $Z_j(x_2^0)$ has zero surface on X (property (a)) and partitions $N_j(x_2^0)$ in a finite number of zones (property (b)). Alternatively, assume that $N_j(x_2^0)$ is the neighbourhood of an isolated root. In this case, by the implicit function theorem, the perturbation defines, for positive ε chosen sufficiently small, a continuous function on $N_j(x_2^0)$ for which $Z_j(x_2^0)$ also has zero surface on X. Therefore, for all j, the set $Z_j(x_2^0)$ has zero surface partitions $N_j(x_2^0)$ in a finite number of zones. This holds in a band around every x_2^0 and since $\varepsilon > 0$, it follows that the unions of all sets $Z_j(x_2)$ has zero surface, and since it holds for an arbitrary pair [s,s′] the property holds for all pair, proving (a) and (b).□

The result has an easy geometric interpretation. For given prices p, the function $z(x_1, x_2)$ can be thought of as a mountain. Consider the intersection of this mountain with a horizontal plane at given altitude (sea level). On this plane, mark the locations where the land that lies below sea level in blue, and those above sea level in black. Because the function is continuously differentiable, the resulting map will now delineate islands and lakes. The distinction will only be ambiguous if the mountain is exactly tangent to sea level. But this tangency will disappear after an arbitrarily small rise in sea level (property (a)). Assumption W(ii) implies that there are no such plains. Property (b) establishes that the number of islands and lakes will be finite.

5. Land consolidation

The zoning approach can also be used to study land consolidation. When applied at village level, the zoning problem (4.3) allows to reassign land among farmers s, and the potential revenue from this process can be evaluated. However, this model has several limitations. One of these is that it does not attach any cost differentials to the configuration of the parcels: their possible fragmentation, imbalanced shape and lack of contiguity do not affect the cost of production. Here we only address the aspects of fragmentation and shape, and neglect the issue of contiguity of parcels, which is highly relevant for problems of real estate management. Nonetheless, a squared Euclidean (or higher) norm will usually create sufficient "gravity forces" to ensure contiguity of most parcels. The first model in this section determines the optimal parcel structure from a combined zoning (assigning fields to households) and location of a facility (in the simplest case: finding the barycentre of the fields cultivated by a farmer). The second model imposes an additional restriction on equitable distribution of gains from consolidation.

In model (4.3), the cost of production only depends on the distance between the point x in the field and the homestead located at b^s. This might be unrealistic, as the farmer can visit his plots more easily if they are located close together than if they are dispersed over a wide area. Hence we must also account for the distance between plots. To obtain such a measure of spread, we allow every farm s to be the user of not more than, say, J distribution centres or roads indexed sj, as in (2.5) above. Farmer s could be thought of as taking his bullock cart with seeds or manure to some roadside location near his plots and distributing these inputs over his fields with manual labour, or conversely, collecting his crops with manual labour before loading them on the cart. We continue referring to b^s as the homestead but it might also the location of the marketing post of farmer s, who could choose to live somewhere along this road.

The main assumption is that there is a finite number of alternative locations given to every farmer (which might be the same for different farmers), and that have the right of passage through neighbouring fields. This ensures that lack of contiguity of parcels does not pose any problem. This amounts to assuming that land fragmentation only matters to the extent it causes parcels to be spread more widely, but the number of parcels itself does not affect cost. Parcels should be close by, but they do not have to be consolidated, and their shape is only relevant to the extent it affects distance.

A model of unconstrained land consolidation

Formally, we assign at most J distribution centres or roads to every farm s. We can, for example, decompose the transportation cost from the field x to the homestead (or market) b^s into two parts: from x to h^{sj} and from h^{sj} to b^s. Thus, total transport costs per unit is a function $w_k^s j(x)$, say,
as in (2.5):

$$w_k^s j(x) = a_k^j \, \|x - h^{sj}\| + b_k^s \, \|h^{sj} - b^s\|, \tag{5.1}$$

where a_k^j is a unit cost for transport, say, by foot, b_k^s a unit cost for transport by cart, and $\|\cdot\|$ denotes a vector-norm and measures the distance between x and b^s, and a distance between centre of operations and the homestead. This yields as an immediate extension of model (4.3)

$$\min_{p^s \geq 0,\ \text{all } s} [\Sigma_s \Pi^s(p^s) + \int \max_{sj} \pi^{sj}(p^s - w^{sj}(x), x)\ dG(x)]. \quad (5.2)$$

It might be added that issues related to the optimal shape of plots can be dealt with along similar lines. Rather than imposing some ideal geometry, one can endow each farmer with more detailed options for establishing central sites. Revenue maximization will then induce a particular concentrated geometry of spots around these sites. Yet the contiguity issue remains unsolved. It would also be interesting to allow for endogenous determination of the points h^{sj} by treating these as endogenous (continuous) variables but this would introduce a nonconvexity in the program.

Regional model with fixed land base

Program (5.2) determines the optimal land consolidation. To represent a given pattern of land fragmentation, one defines the distribution $G^s(x)$ which specifies the given association between point x and farmer s.

$$\min_{p^s \geq 0,\ \text{all } s} [\Sigma_s \Pi^s(p^s) + \Sigma_s \int \max_j \pi^{sj}(p^s - w^s(x), x)\ dG(x)], \quad (5.3)$$

This is actually a regional model with separate farms and a fixed land base per farm. Its solution will, for every farm, describe the optimal land allocation to various crops and the zones indexed j will become the fields of the farm. Moreover, every regional model will have a differentiable supply response function $q^s(p^s)$, even if we relax the strict convexity in assumption r, and allow for every point to choose not only the best road j but also the most profitable crop k, with full specialisation i.e. solve the "linear program": $\max_{jk} \pi^{sjk}(p^s - w^s(x), x)$.

Restrictions on land consolidation

Land consolidation transactions obviously generate a change in the land distribution. There is a distinction between a centralized and a decentralized process of transactions. In a centralized land consolidation process all participants start the process by pooling their land resources together. Next the optimum of (5.2) is computed, jointly with the cost for every site (household). Finally, a compensation is paid to the losers, which is financed either by taxing those who gain, or from external funds (village development), according to some agreed rules for sharing the surplus from cooperation. In many situations, the external funds to compensate losers directly or through development of infrastructure will be lacking. In addition, participants are often reluctant to accept compensation payments, because the consolidation process has lasting consequences and calls for recurrent payments which, in view of weak enforcement mechanisms, might not be forthcoming once the consolidation process has been completed. Moreover, farmers will rarely be able to provide side payments in cash.

Under such conditions, participants will tend to prefer a system that guarantees to some fairly distributed gain in the direct revenue from farming, and maximizes the relative gain of the least-favoured. It is relatively straightforward to modify (5.4) accordingly. Let V_0^s denote the original income of farmer s in (5.4), which is assumed to be positive. We will solve:

$$\max_{\substack{q^s \geq 0,\ y_k^{sj}(x) \geq 0 \text{ all } k,\, s, \text{ all } x}} \min_s \, [1/V_0^s[R^s(q^s) + \int (r(-\Sigma_{sj}\, y^{sj}(x), x) - \Sigma_{sj}\Sigma_k\, w^{sj}(x)\, y_k^{sj}(x))\, dG(x)]$$

subject to (5.4)

$$q_k^s - \Sigma_j \int y_k^{sj}(x)\, dG(x) \leq 0, \text{ all } k \text{ and } s$$

The structure of this problem differs from that of the earlier models because of the $\min_s$ operator. This requires some reformulation.

Solution procedure

Let us define the S-dimensional simplex $\Lambda = \{(\lambda^1, ..., \lambda^S) \geq 0 \mid \Sigma_s \lambda^s = 1\}$ and represent the $\min_s \int f^s(\cdot)\, dG(x)$ part as $\min_{\lambda\in\Lambda} \int \Sigma_s \lambda^s f^s(\cdot)\, dG(x)$, because this enables us to compute the gradients directly. The variables λ^s can be interpreted as welfare weights (or as value-added tax rates incremented by unity). We also define scaled prices $\tilde{p}^s = \lambda^s p_s$, and scaled the profit functions terms

$$\tilde{\Pi}^s(\tilde{p}^s, \lambda^s) = \max_{q^s\geq 0} [\lambda^s R^s(q^s) - \Sigma_k\, \tilde{p}_k^S q_k^s], \qquad \text{all } s \tag{5.5a}$$

which is convex non-increasing in scaled input price $\tilde{p}^s$ and convex non-decreasing in welfare weight λ^s. For spot x, the scaled profit function is:

$$\tilde{\pi}^{sj}(\tilde{p}^s - \lambda^s w^s(x), x) = \max_{y^s(x)\geq 0} [\Sigma_k\, (\tilde{p}_k^S - \lambda^s w_k^{sj}(x)) y_k^s(x) + \lambda^s r(-y^s, x)], \tag{5.5b}$$

Noticing that $\tilde{p} \in \mathcal{P}$ since $0 \leq \tilde{p} \leq p$, the resulting nested formulation is:

$$\min_{\tilde{p}\in\mathcal{P},\lambda\in\Lambda} \int v(\tilde{p}, \lambda, x)\, dG(x), \tag{5.6a}$$

for

$$v(\tilde{p}, \lambda, x) = 1/V_0^s\, [\tilde{\Pi}^s(\tilde{p}^s, \lambda^s) + \max_{sj} \tilde{\pi}^{sj}(\tilde{p}^s - \lambda^s w^{sj}(x), x)], \tag{5.6b}$$

which is jointly convex in $(\tilde{p}, \lambda)$. We can solve this problem by an SQG-procedure similar to (4.9): at each step t = 1, 2, ... the price vector $\tilde{p}(t)$ is adjusted in the direction $\partial v(\tilde{p}(t), \lambda(t), x(t))/\partial \tilde{p}$ and $\lambda(t)$ in the direction $\partial v(\tilde{p}(t), \lambda(t), x(t))/\partial\lambda$. We propose the following decentralized stochastic adjustment procedure:

$$\tilde{p}(t+1) = \Pi_{\mathcal{P}}(\tilde{p}(t) - \rho_t\, \partial v(\tilde{p}(t), \lambda(t), x(t))/\partial\tilde{p}), \tag{5.7a}$$

$$\lambda(t+1) = \Pi_{\Lambda}(\lambda(t) - \rho_t\, \partial v(\tilde{p}(t), \lambda(t), x(t))/\partial\lambda), \qquad t = 1, 2, ... \tag{5.7b}$$

where ρ_t is the step-size that is supposed to satisfy (2.3). Note that, as in (4.9), the price adjustment is a stochastic Walrasian tatonnement, but in addition, the weights are adjusted so as to give lower weight to those who stand to gain most, relative to the original situation V_0^s. We summarize this as a proposition.

Proposition 3 (Decentralised land consolidation without losers): Let assumptions R, r and W hold. Then, with probability 1, procedure (5.7) with stepsize ρ_t satisfying (2.3) converges to the (global) optimum of (5.6) which also solves (5.4).

Proof. Since $v(\tilde{p}, \lambda, x)$ is jointly convex in $(\tilde{p}, \lambda)$ on the compact, convex domain $\mathcal{P}\times\Lambda$, convergence follows as in section 2.□

Procedure (5.6) is remarkably simple and transparent, despite the apparent complexity of problem (5.4). We start from given $\tilde{p}(t)$ and $\lambda(t)$ select a point $x(t)$ at random. Next, we choose the most profitable destination sj, for all commodities produced at x to be shipped to. Finally, we compute the SQG as derivative w.r.t. $\tilde{p}^s$ and λ^s of the associated "total profit" function $1/V_0^s\ [\tilde{\Pi}^s(\tilde{p}^s, \lambda^s) + \tilde{\pi}^{sj}(\tilde{p}s - \lambda^s w^{sj}(x), x)]$ (if there are several, choose, say, the one with lowest index value).

6. Conclusion

In this paper, land was treated as a perfectly divisible commodity and this enabled us to reach relatively strong conclusions about the convergence of the processes of land allocation and consolidation, even when restrictions are being imposed on transactions. However, in reality land transactions are difficult and the question is, therefore, whether the models and adjustment processes presented here bypass important elements.

First, land in reality is indivisible to some extent. Farmers own plots of different size and quality. and transactions relate to plot surfaces, not to points, precisely because agents cannot keep on transacting forever. Now, if the surface X is partitioned into N fixed parcels the optimal land consolidation model (5.6) becomes

$$\min_{\tilde{p}\in\mathcal{P},\lambda\in\Lambda} V^N(\tilde{p}, \lambda) \tag{6.1a}$$

for

$$V^N(\tilde{p}, \lambda) = 1/N \sum_{i=1}^N [\Sigma_s\ 1/V_0^s\ [\tilde{\Pi}^s(\tilde{p}^s, \lambda^s) + \max_{sj} \tilde{\pi}^{sj}(\tilde{p}^s - \lambda^s w^{sj}(x^i), x^i)] \tag{6.1b}$$

To solve this problem, i.e. to clear the land market, the decentralized SQG-procedure (4.9) can be used as before, now with $x(t)$ sampled from the discrete distribution $\{x^1, ..., x^N\}$ instead of $G(x)$. The difference between the calculated optima $V^N(\tilde{p}, \lambda)$ and $V(\tilde{p}', \lambda)$ will estimate the welfare loss from maintaining a given parcel size as well as associated price distortion $(\tilde{p}-\tilde{p}')$. A similar discretization approach can be used for the other models in this paper. Clearly, this presupposes that the SQG-procedure will find the global minimum of (6.1), which is by no means assured because the profit functions $\tilde{\pi}^{sj}$ defined over the parcel might differ from those at say the mid point of the parcel and therefore if these are very heterogeneous, not be convex in prices $\tilde{p}$.

Secondly, the resulting optimum would often be significantly lower than the unconstrained one, even for divisible parcels, and this might frustrate the progress because some potential gainers might prefer the status quo to a time consuming negotiation process that only brings marginal improvement, and might even end before a Pareto-efficient solution is reached, i.e. as soon as the least favoured cannot improve their condition.

Finally, process (5.7b) illustrates that in the course of the decentralized adjustment process, the welfare weights λ^s are not necessarily zero for all but the least favoured. Hence, in the course of the process their will as a rule be land allocations that yield income losses for some agents and gains for others. Then, both in both groups there might be agents who want to step out of the process: the losers because they are afraid to lose more and the winners because they want to consolidate their gains. Thus the process of transactions might end for lack of participants, long before the optimum has been reached. It is possible to adjust the SQG-procedure to account for this, by restricting the random selection of points to an appropriately defined part of the distribution. One possibility is to look for "matching" pairs of transactions. This raises various questions about the conditions under which determine whether the process will converge to a Pareto-optimum or stagnate. This is a subject for further investigation.

Another challenge for further research is to characterize the type of non-convexities and discontinuities that can be "healed" by the integral operation. Lemma 1 has illustrated the advantage of modelling in a spatial continuum, as opposed to a discretized space, since this allows to neglect specific discontinuities and non-convexities, as characterized by the "almost everywhere"-qualifier. However, we have only considered problems for which this qualification is given as part of the problem specification. It would seem useful to formulate conditions on the model itself which allow to neglect non-convexities and discontinuities that appear in it, because of the "healing" effect of integration. In the model of pollution control this might enable us to study increasing returns in pollution, and in the land consolidation model to determine endogenously the location of distribution centres.

References

Avriel, M. (1976) *Nonlinear Programming*. Englewood Cliffs, NJ: Prentice Hall.

Beck, R.L, and J.D. Goodin (1982) 'Optimum number and location of manufacturing milk plants in Kentucky', University of Kentucky, College of agriculture, Bulletin 718.

Drezner, Z., (ed.) (1995) *Facility location analysis: survey of applications and methods*. New York: Springer.

Ermoliev Yu.M. (1988), 'Stochastic quasigradient methods'. in Yu. Ermoliev R.J.B. Wets, eds. *Numerical Techniques for stochastic optimization*, Numerical Math,: Springer-Verlag.

Ermoliev, Yu., M.A. Keyzer, V. Norkin (1997) 'Global convergence of Walrasian tatonnement', Centre for World Food Studies, Amsterdam.

Ginsburgh, V., and M.A. Keyzer (1997) *The structure of applied general equilibrium models.* Cambridge, MA.: MIT-Press.

Hansen, P. Labbé, M., Peeters, D., and J.-F. Thisse (1987) 'Facility Location Analysis', *Fundamentals of Pure and Applied Economics*, 22 :1-70.

Ortega, J.M., and W.C. Rheinboldt (1970) *Iterative solution of nonlinear equations in several variables*. New York: Academic Press.

Rust, J. (1997) 'Randomization to break the curse of dimensionality', *Econometrica*: 65:487-516.