THE AUSTRALIAN JOURNAL OF AGRICULTURAL ECONOMICS

VOL. 30 APRIL 1986 NO. 1

PLANT LOCATION ANALYSIS USING DISCRETE STOCHASTIC PROGRAMMING*

COLIN G. BROWN and ROSS G. DRYNAN†
University of Queensland, St Lucia, Qld 4067

A plant location model with two major aspects is outlined. First, discrete stochastic programming is used to handle variability in supplies and demands. Second, the cost structure of plants is modelled in more detail and with more realism than is usual. Results from applying the model to the Queensland cattle slaughtering industry demonstrate the inappropriateness of using traditional deterministic plant location models to analyse problems with major stochastic elements. Deterministic models yield plant locations, sizes, throughputs, commodity flows and implications which differ markedly from those generated by stochastic models in which plant sizes and locations are optimally matched to variable fat cattle supplies. In addition, the traditional deterministic long-run model overestimates the normative gains of industry rationalisation.

Since King and Logan's (1964) seminal work, the methodology of studying plant location has developed progressively to account for multiple products (Hurt and Tramel 1965), imperfect competition (Ryland and Guise 1975) and plant dynamics (Kilmer, Spreen and Tilley 1983). A further extension of plant location analysis, namely the inclusion of variability, is described in this paper. Such an extension is worth examining in the Australian context, where variability of supply and demand has been strongly evident for most rural commodities (Anderson 1979).

While techniques which accommodate variability have been developed in most areas of agricultural economics, there has been little attention afforded it in the plant location literature. By ignoring variability, traditional plant location models rule out a priori the possibility of optimum plant location, size and commodity flows being sensitive to variability in supplies and demands. Further, traditional models overlook the fact that variation in plant throughputs, *per se*, can create significant costs. And these are not restricted to physical costs. Variation in throughputs is likely to result in social costs associated with labour displacement and its damaging effects on local economies based primarily on the operation of a particular plant (Industries Assistance Commission 1983). Optimal plant location patterns developed ignoring

* An earlier version of this paper was presented at the 29th Annual Conference of the Australian Agricultural Economics Society, University of New England, Armidale, 12–14 February 1985.

† Colin Brown is now at the Bureau of Agricultural Economics. The authors are grateful for comments from John Longworth and *Journal* referees.

variability need not be optimal, or even attractive, in terms of social well-being.

The specific approach for handling variability is outlined in the next section. It entails the marrying of well-known discrete stochastic programming methods with traditional plant location models. As with most plant location analyses, total cost is adopted as the minimand. While a major empirical objective was to apply the model to studying the location of abattoirs in the Queensland cattle slaughtering industry, another was to make an assessment of the impact of modelling variability in that application. This latter empirical analysis is reported in this paper.

The Discrete Stochastic Plant Location Model

The traditional plant location problem seeks to minimise the sum of assembly, processing and distribution costs by selecting plant sites, sizes, throughputs and product flows subject to the available supplies and demands. It takes the following form:

(1) Minimise $\Sigma_{tp} A_{tp} X_{tp} + \Sigma_p f_p(X_p) + \Sigma_{pj} D_{pj} X_{pj}$

subject to:

(2) $\Sigma_p X^S{}_{tp} \leq Q^S{}_t$ for all t

(3) $\Sigma_p X^D{}_{pj} \geq Q^D{}_j$ for all j

(4) $\Sigma_t X^S{}_{tp} = \Sigma_j X^D{}_{pj} = X_p$ for all p

(5) $X^S{}_{tp}$, X_p and $X^D{}_{pj} \geq 0$ for all t, p and j

where A_{tp} = unit assembly costs from supply region t to plant p;
$f_p(X_p)$ = total processing costs of plant p;
D_{pj} = unit distribution costs from plant p to demand region j;
$X^S{}_{tp}$ = number of units of raw product shipped from supply region t to plant p;
X_p = number of units of raw product processed at plant p;
$X^D{}_{pj}$ = number of units of processed product shipped from plant p to demand region j;
$Q^S{}_t$ = supplies in region t; and
$Q^D{}_j$ = demands in region j.

In practice, specification of the processing costs has ranged from that of a constant average variable processing cost for all throughputs, through that of a linearised cost function with respect to throughput, to a more general usage of a linearised cost function in which integer variables are used to account for economies of size in processing. The specification depends, in part, on which type of analysis is involved. Two types are possible. Optimal plant locations, sizes and commodity flows can be determined in a long-run analysis in which processing costs ($f_p(x_p)$) involve both capital and variable costs. (Capital costs are here considered as the amortised costs of plant investment and other ownership costs. All other plant processing costs are viewed as variable costs.) Optimal commodity flow patterns alone can be identified in a short-run (trans-shipment problem) analysis, in which plant locations and sizes are fixed, by specifying $f_p(X_p)$ as variable processing costs only.

One of the shortcomings of the traditional approach, particularly for long-run analysis, is its failure to recognise variability in the raw product supplies and the demands for processed goods. Variability in the constraints of programming models can be handled either as continuous variation or as discrete variation. In the former, and arguably more realistic case, computational difficulties limit the extent to which the reactions to a particular scenario can be modelled. The method of two stage linear programming with recourse is an example of this approach (Charnes, Cooper and Thompson 1965). Balachandran and Jain (1976) used it in their investigation of optimal plant location under random demand by including stochastic constraints (replacing equation 3 and assigning penalty costs to surplus or slack demands (in the objective row, equation 1. But this procedure is less than ideal. It does not explicitly allow for the real options (and their costs) for meeting shortages and disposing of surpluses when they arise. No information is provided on the optimal commodity flow solutions for different demand (or supply) situations.

The second approach to handling constraint variability involves discrete stochastic programming (for example, Cocks 1968). To the authors' knowledge, there has not been an attempt to use this approach in plant location models. An elementary discrete stochastic programming problem has the following form:

(6) Minimise $C'_d X_d + \Sigma_s w_s C'_s X_s$

subject to:

(7) $A_{dd} X_d \leq B_d$

(8) $A_{ds} X_d + A_{ss} X_s \leq B_s$ for $s = 1, \ldots, n$

(9) $X_d, X_s \geq 0$

where C_d, C_s, X_d, X_s, B_d and B_s are vectors; A_{dd} and A_{ds} are conformable matrices; n is the number of states of nature; and w_s is the probability of occurrence of state of nature s.

Activities and constraints are grouped into $n+1$ subsets. Each of the last n subsets (subscripted by s) refers to a particular state of nature. The activities in a set are performed only in the corresponding state of nature and must satisfy the set of constraints applying for that state of nature. Their contributions to the objective function are weighted by the probability of the state of nature occurring. The first set of activities (subscripted by d) lies at the heart of discrete stochastic programming, for without it, activity levels for each state of nature could be chosen independently in a deterministic planning problem. The activities in this set represent those that are selected and undertaken irrespective of the state of nature. These 'common' activities have implications for the objective function and for the feasibility of activities specific to a state of nature. Their optimal levels depend very much on their interactions with the state-specific activities.

In essence, this simple discrete stochastic programming formulation implies a two-period decision problem. Some decisions must be made and implemented, and subsequently other decisions are possible once details of the state of the world become known.

The plant location problem has precisely this sequential form. The decisions about locations and sizes of plants, once implemented, hold for all supply and demand scenarios that may eventuate.[1] These decisions are modelled by the common activities of the stochastic program. Subsequent decisions are made regarding sources of supplies, processing and distribution once the supply–demand scenario becomes known. These decisions can be reflected in a group of activities defined for the particular scenario.

The advantage of modelling variability through a discrete stochastic programming formulation is that it allows for specification of detailed adjustment options and costs, and the long-run decisions of optimal plant size and location are forced to take account of all possible states of nature and available adjustment options. Its limitations lie in the realism of a limited number of states of nature. However, in the presence of the law of diminishing marginal returns to modelling efforts, variation in plant location problems may more efficiently be categorised by a relatively small number of discrete states of nature than by the infinity implied by assuming continuous variation.

The core of the discrete stochastic plant location model is presented in algebraic form in the Appendix. The Appendix demonstrates that the model combines the elements of the basic plant location and stochastic programming models outlined above. But there are added complexities. Consideration is now given to some conceptual aspects of the model. How these conceptual aspects are incorporated into the programming model is briefly described in the Appendix. More detail is available in Brown (1985).

Plant capacity

The notion of capacity here is one of the size of the plant. But capacity does not represent any absolute limit on the processing ability of the plant. Rather it is an index, albeit measured in the same units as throughputs. A plant of a given capacity will entail some particular capital costs, and will have associated operating characteristics reflected in a curve describing the costs of using a plant to process various throughput levels, both above and below capacity. As will become clear later, in choosing our index, we have followed de Leeuw (1962) and others who define capacity as that throughput which results in minimum average variable processing costs.

Cost structure

A major feature of the formulation is the detailed specification of long-run costs. The traditional long-run cost curve used in plant

[1] The subjective probabilities of occurrence of various supply and demand scenarios are assumed not to change over the planning horizon. To allow for such changes, the model would have to allow for subsequent decisions on location and sizes of plants.

location studies relates total cost to capacity alone. But this curve is inappropriate as a planning concept in the presence of variability because costs depend on both capacity and throughput.

Long-run processing costs are modelled by a capital processing cost curve (a function of capacity) and separate variable cost curves defined for each possible size of plant (a function of throughput and the given plant capacity). A capital processing cost curve for a range of plant sizes is presented in Figure 1(a). Variable processing cost curves for two different size plants (A_1 and A_2 of Figure 1(a)) are specified in Figure 1(b). (The curves shown in Figures 1 and 2 are hypothetical and used for pedagogic reasons alone. More complex shapes are possible.) At low throughputs, plant A_1 has lower variable processing costs than A_2, reflecting, for example, larger lump sum costs for the larger plant A_2. These costs relate to such things as cleaning, refrigeration and lighting. However, at some throughput (beyond the capacity of A_1) the variable processing costs of A_1 will rise to exceed the variable processing costs of A_2.

Although plants of various sizes are likely to have different variable processing costs for any given throughput, in a modelling framework it is computationally infeasible to be completely general. To maintain linearity, it is necessary either to limit plant capacity to one of a set of discrete sizes, or to make more assumptions about the behaviour of variable processing costs with changes in plant size. The latter option was selected. In particular, the variable processing cost curve defined for a particular plant size is assumed to be a radial projection of the variable processing cost curves defined for other size plants.

Although A_1 and A_2 have different variable processing cost curves in Figure 1(b), they have been drawn to reflect the particular cost structure. For example, consider the 100 per cent capacity utilisation points on A_1 and A_2 (points X and Z in Figure 1(b), indicating throughputs of 100 and 200, respectively). These points are constructed to lie along the same ray. That is, they have the same average variable processing costs ($10). Similarly, other points of equal capacity utilisation exhibit the same average variable processing costs. For example, consider the points W and Y on the variable processing cost curves indicating capacity utilisations of 25 per cent (or throughputs of 25 and 50 for A_1 and A_2, respectively). These points again lie on the same ray and so have the same average variable processing cost (namely $30). The variable processing cost curve for A_2 is simply a scaled up version or radial projection of the variable processing cost curve for A_1. The basic assumption, therefore, is that plants operating at the same capacity utilisation, no matter what their size, have the same average variable processing cost.

This enables the definition of a single variable processing cost curve to represent plants of all sizes (see Figure 2a). Here variable processing costs are defined per unit of capacity, with the horizontal axis now measuring capacity utilisation and not throughput. Calculation of the variable processing costs associated with a particular throughput in a particular plant involves multiplication of the curve A by the capacity of a plant. For example, from Figure 2(a) variable processing costs per unit capacity are listed as $10 for 100 per cent capacity utilisation. Therefore, for plants A_1 (size = 100) and A_2 (size = 200), variable processing costs associated with 100 per cent capacity utilisation (throughputs of 100 and 200, respectively) will be \$1000 (100×\$10) and \$2000 (200×\$10), respectively. These are the same as the variable processing costs for points X and Z (throughputs of 100 and 200, respectively) as shown in Figure 1(b). Hence, information contained in curve A of Figure 2(a) is sufficient to specify the entire variable processing cost curves for plants A_1 and A_2 in Figure 1(b) or indeed the variable processing cost curve for any plant irrespective of its size.

It may be unrealistic to assume that plants will exhibit the same average variable processing cost behaviour as a function of capacity utilisation irrespective of size. Plants based on different technologies could be expected to have different operating characteristics. It may also be expected that large plants will have lower average variable processing costs at 100 per cent capacity utilisation than smaller plants operating at the same capacity utilisation. In this case, a number of capacity zones may be defined as shown in Figure 1a. A unique variable processing cost curve can be specified for each zone (see Figure 2b). For example, in relation to cattle processing, zone A may refer to small bed-dressing operations, zone B to medium on-the-rail slaughtering systems, and zone C to large technologically advanced plants. All plants within a particular zone will exhibit the same relationship between average variable processing costs and utilisation of capacity.

Multiple objectives

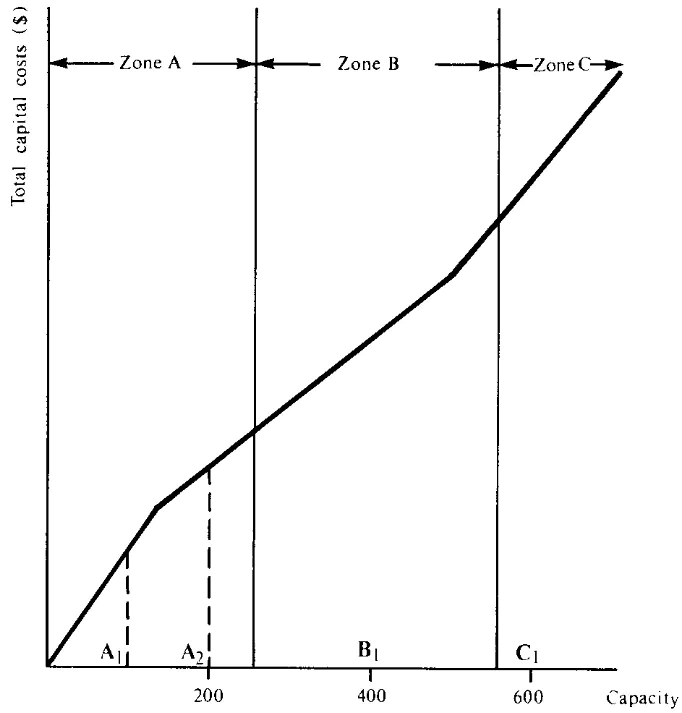
Rarely do real world planning problems involve a clear, single-dimensioned objective. Because the stochastic formulation focuses on each state of nature and calculates product flows and throughputs for each, account can be taken of objectives associated with variation (for example, stability of employment and throughputs).

The variation in throughputs at a plant can be measured by methods similar to the MOTAD formulation espoused by Hazell (1971). That is, constraints and activities defining the deviations of throughputs around the mean for a plant can be added to the matrix. It remains, then, to define the objectives and/or the costs associated with the measured deviations in throughput. The costs probably depend on the directions of changes in throughputs and are likely to be non-linear. When plant closure occurs, they may depend on the length of closure and the level of throughputs prior to and after the closure. Undoubtedly a quite detailed formulation of these complexities is possible, even in the context of a programming model, if the relevant supporting data are available.

A much simpler approach has been used here. The costs are assumed to be directly proportional to the mean absolute deviation of plant throughput. If this were considered too simplistic, constraints on absolute capacity and mean throughput could be added. If desired, one could then force a plant to close unless it operates at sufficiently high capacity utilisation levels and/or with throughputs close to its mean throughput level.

Plant Location Matrix Generator

As part of an attempt to control the costs of plant location analyses, a general plant location matrix generator (PLOMAG) was developed to



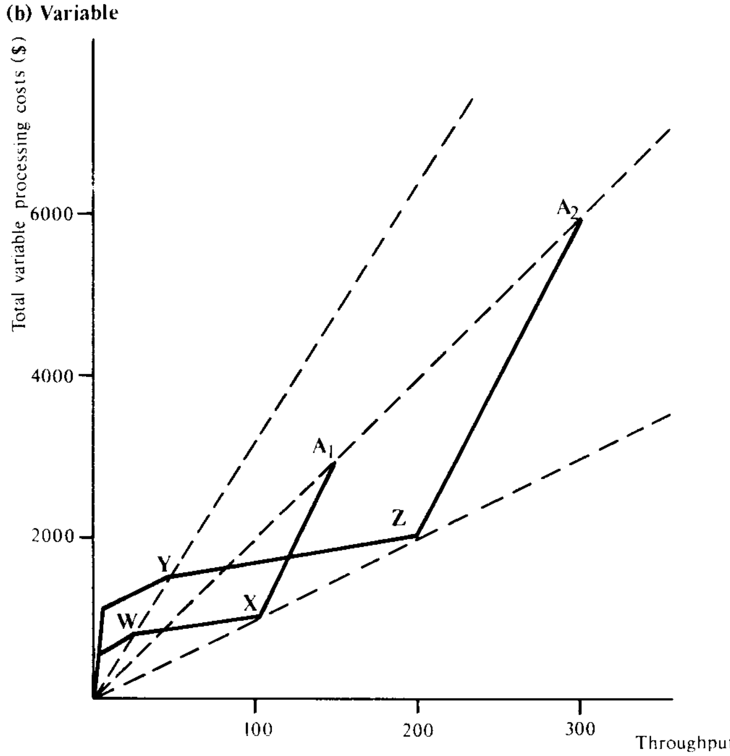


FIGURE 1—Total Processing Cost Curves.

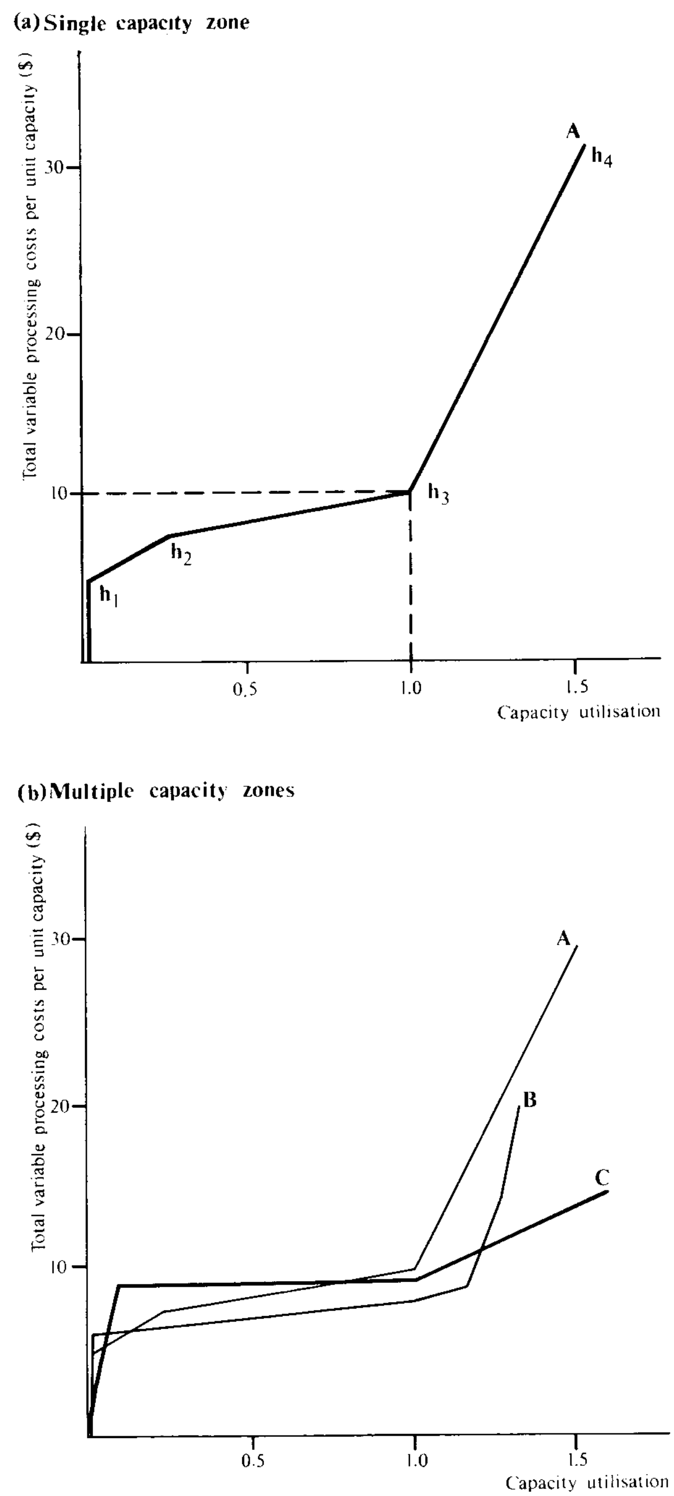


FIGURE 2—Total Variable Processing Costs as a Function of Capacity Utilisation.

enable rapid formulation of stochastic (and deterministic, if desired) plant location problems. PLOMAG accepts as input raw data such as supplies, demands and costs, first transforming them into the programming matrix discussed above. Another routine converts this matrix into the MPS input format as required by commercially available mixed integer programming packages,[2] such as APEX III (Control Data Corporation 1979).[3]

PLOMAG generates the programming matrix for all the variability-related concepts discussed above. Other formulations, including constraints on throughputs and capacities, restrictions on particular shipping activities, modification of integer permission activities, average throughput and capacity restrictions, and maximin decision rules for multiple objectives analysis are also handled (Brown 1985).

Abattoir Location for the Queensland Cattle Industry

Following a buoyant period in the late 1970s, the Australian meat processing industry took a downturn in the early 1980s. Many processing firms either closed their abattoirs or operated at low utilisation levels. A government inquiry was initiated into the Australian abattoir and meat processing industry (Industries Assistance Commission 1983) and a variety of industry rationalisation proposals was put forward. This was the context in which the empirical work was undertaken.

Around two million head of cattle are slaughtered annually in Queensland. The geographical distribution of supplies, demands and processing centres means large shipping distances and significant transport costs. In contrast to supplies, which are distributed throughout the entire state, processing is concentrated in the south-east and demand is centred on Brisbane. Around 80 per cent of the meat produced is exported, chiefly through Brisbane. In 1984, there were 39 abattoirs operating at throughput levels ranging from less than 100 to more than 1000 head of cattle per day.

Large year-to-year variations occur in slaughter cattle supplies because of a variety of physical and market uncertainties. In addition, a marked seasonal variation (particularly in north Queensland) is exhibited as a result of factors such as feed quality and availability, and the ability of producers to move stock. Supplies peak during the period from March to October.

The basic model used in the analysis has eighteen supply regions, four demand regions and twenty potential plant sites. Variation in supplies was characterised by: (a) 'between-year' variation described by three states of nature, namely 'good', 'average' and 'poor' cattle producing years with subjective probabilities of occurrence of 0.2, 0.55 and 0.25, respectively; and (b) seasonal variation described by two seasons, 'peak' and 'off-peak', of 160 and 80 working days, respectively.[4]

[2] The model uses integer activities for three purposes: handling economies of capacity, scale economies of throughput and multiple capacity zones. Problems exhibiting none of these features reduce to ordinary linear programs.

[3] However, as the programming matrix is always developed, other mixed integer programming packages would be compatible with PLOMAG by slight modification of the second conversion routine.

[4] Details of how the variation was described and basic data for the empirical analysis can be obtained from the authors.

Variability and short-run solutions

The effects of variability were initially examined with plant capacities fixed at existing levels. Two runs were made for each of the three types of years, the runs differing in that one (seasonal model) recognised the seasonal variation with peak and off-peak periods, while the other (average model) consisted of a single period of 240 working days. Although plants could not reduce or increase capacity within the annual time horizon, they could close down in a particular season or otherwise alter throughput levels. The results of these six runs are reported in Table 1.

In the poor year, the solutions of the average and seasonal models are quite different. The main effect of variability in the poor year appears to be as follows. In the average model, daily supplies in north Queensland are too low to maintain high capacity utilisations (and low processing costs) in the larger plants, and they close. By contrast, in the seasonal model, the larger plants can attract sufficient daily supplies to operate at high capacity utilisation in the peak season, but are forced to close when faced with the reduced supplies of the off-peak season. Other Queensland plants, such as the large plants in the Moreton region and those at Toowoomba and Biloela, also exhibit considerable differences between the runs. This reflects the variation in capacity utilisation between the peak and off-peak seasons in the seasonal model.

Introducing variability has different effects on the results in poor and average years. In the poor year, ignoring seasonal variation reduces north Queensland processing and increases that in central Queensland. But in the average year, the importance of north Queensland is greater when seasonality is recognised. The average model favours north Queensland in the average year because average daily supplies are now sufficient for plants to remain open and operate at relatively high utilisation. That is, north Queensland plants do not have to close or operate at low utilisations in the off-peak season. The local off-peak supplies that north Queensland plants process in the average model go to south Queensland plants in the seasonal model.

In the good year, variability again has implications, but for reasons different to those applying in the average and poor years. The effects are reversed. When variability is ignored, south Queensland plants are favoured at the expense of north Queensland plants. Supply shipment patterns indicate that the difference is associated with overcapacity processing. In the seasonal model for the good year, peak season supplies are so large that all plants operate at full utilisation. North Queensland plants then process not only all supplies within their region but also significant central Queensland supplies. This movement of supplies is against the normal flow of supplies toward Brisbane but arises because the increase in transportation costs is less than the increase in processing costs resulting from operating at overcapacity levels. When variability is ignored in the good year, daily supplies are low enough for all of the central Queensland supplies which are processed by north Queensland plants in the seasonal model and for some north Queensland supplies to be diverted to the larger south Queensland plants without costly overcapacity processing.

These runs identified two major aspects of the effect of variability on short-run plant location analysis. The first is the significant impact on throughputs, capacity utilisations and shipment patterns. The average models arrive at values which differ markedly from the optimal values obtained from the seasonal model. Capacity utilisations suggested by the average model simply do not apply in both the peak and off-peak seasons and, in most cases, would be viewed by the industry as irrelevant given the reality of seasonal variation. This problem is compounded when, in traditional plant location studies, a single deterministic run ignoring both seasonal and between-year variation is made.

The second aspect highlighted by these empirical results is that the effect of variability is itself variable. It is unlikely to be known a priori. Hence, although some general trends associated with variability may be observed, a model incorporating variability is required when variation in supplies and/or demands arises.

Variability-cost trade-offs

The trade-off frontier between plant throughput variation costs and combined shipping and processing costs (calculated here only for the average state of nature) is shown in Figure 3. That is, it shows the

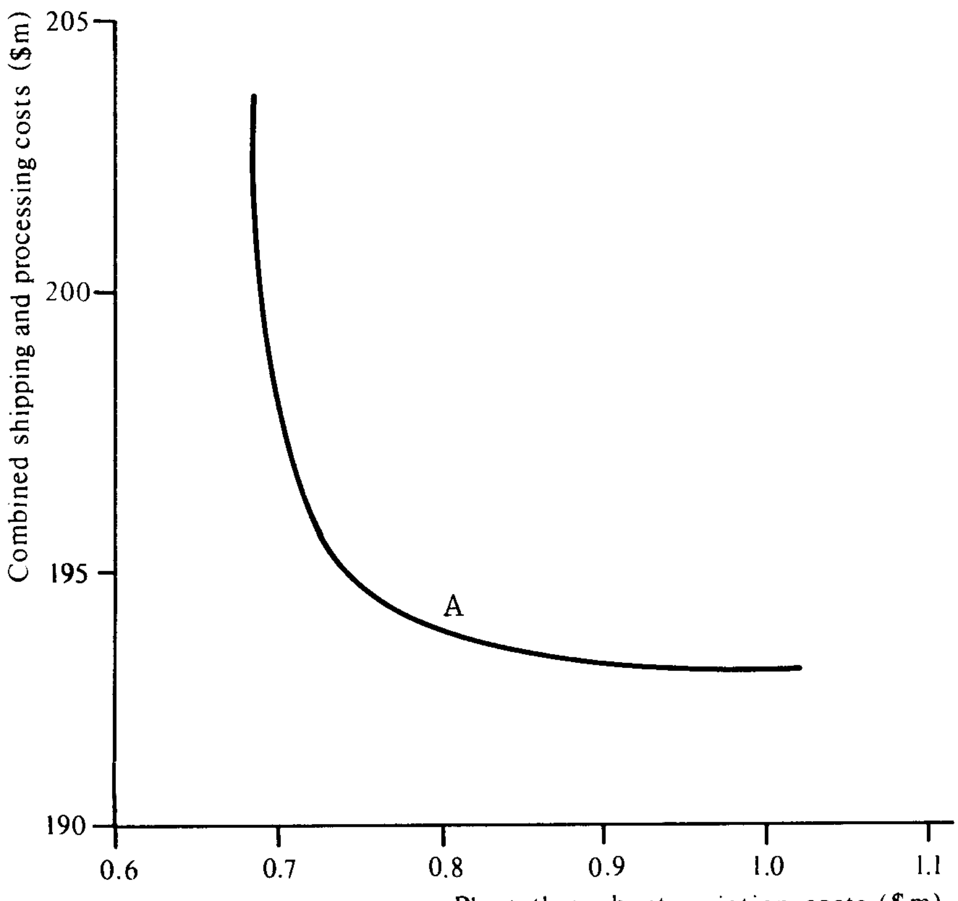


FIGURE 3—Trade-off Frontier of Combined Shipping and Processing Costs versus Plant Throughput Variation Costs.

minimum plant throughput variation costs for any given level of shipping and processing costs. Estimates of plant throughput variation costs were based on a number of factors including size of plant, population size of region (or alternative employment prospects) and prior expectation of variability. The shape of this trade-off frontier is important: there is scope for a non-trivial trade-off of costs. Provided the machinery existed for intervention, it would be possible to obtain a solution which promised large reductions in throughput variation costs with only slight increases in processing and shipping costs. One such solution is shown as point A in Figure 3, with specific solution details presented in Table 1 (Frontier point). On the other hand, a policy which placed excessive emphasis on minimising either plant throughput variation costs or processing and shipping costs would incur significantly increased total costs.

Variability and long-run solutions

Solutions were obtained from four 'clean-slate' or long-run models. These solutions are shown in Table 2. The full clean-slate model incorporates the three types of years and the two seasons. The others involved some simplification. The first (deterministic model) incorporates neither source of variability. It parallels the traditional long-run plant location model. The second incorporates the between-year variability, but no seasonal variability. The third incorporates seasonal variability, but no between-year variability.[5]

The capacity determined in the deterministic model was approximately 2000 head per day less than that determined in the full stochastic model. Throughputs (which are equal to capacity in this single state of nature model) are markedly different from those in the full stochastic model.

Taking account of between-year variation alone leads to total capacity in excess of that in the full stochastic model. In the full stochastic run, capacities were aligned with throughputs in the average year peak season, taking into account both over and under utilisation processing costs, particularly the extreme high and low throughputs in the good peak and poor off-peak seasons, respectively. In the between-year variation model, because the low poor off-peak throughputs did not have to be accounted for, the trade-off between capital costs and under utilisation processing costs on the one hand with over utilisation costs on the other meant that capacities were aligned to avoid over-capacity processing. That is, capacities were aligned with throughputs in the good year. Capacity utilisations and throughputs for all plants in the solution were again markedly different from throughputs in all seasons and types of years in the full stochastic model.

[5] In contrast to all other runs, the between-year variation and full stochastic model runs were terminated short of knowing that a global optimum had been reached. However, the branch and bound mixed integer programming algorithm gives a progressively better lower bound on the optimal least-cost objective value. Each final solution for these runs was no more than 0.1 per cent above this bound, and hence within 0.1 per cent of the optimum. Typically, runs involving a matrix size of 900 column activities, 200 row constraints and 20 integer variables took 20 CPU seconds to reach the global optimum using the Control Data Corporation's APEX III mixed integer programming package on the CSIRONET system.

TABLE 1

Capacity Utilisations (percentages) for the Short-Run Models

		Poor year			Average year			Good year			Frontier point	
		Seasonal		Average	Seasonal		Average	Seasonal		Average		
Region	Capacity[a]	Peak	Off-peak		Peak	Off-peak		Peak	Off-peak		Peak	Off-peak
South Queensland	G	75.4	39.2	62.0	92.9	52.8	76.2	107.4	63.2	99.9	96.0	61.6
Brisbane	E	61.2	65.6	59.2	86.2	59.3	64.4	100.0	61.2	100.0	100.0	57.4
Moreton	F	82.8	22.8	60.8	100.0	60.2	81.8	115.0	57.3	100.0	99.1	60.0
Toowoomba	C	72.3	22.3	59.5	85.5	22.3	72.3	100.0	69.1	100.0	85.5	85.5
Warwick	A	100.0	100.0	100.0	100.0	0.0	100.0	100.0	100.0	100.0	100.0	100.0
Burnett	B	66.7	59.6	64.3	77.6	68.9	74.7	100.0	79.7	86.7	77.6	16.7
Central Queensland	E	62.3	62.1	72.8	79.7	64.8	72.9	100.0	69.4	82.4	76.5	33.5
Biloela	A	0.0	0.0	100.0	100.0	0.0	100.0	100.0	100.0	100.0	100.0	100.0
Rockhampton	D	59.6	59.3	59.1	69.5	63.4	59.3	100.0	54.0	89.1	64.6	0.0
Mackay	B	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
North Queensland	F	42.5	4.2	26.5	64.3	4.2	51.5	100.0	19.0	59.2	58.5	4.2
Bowen	B	80.8	0.0	0.0	78.0	0.0	99.6	100.0	100.0	77.2	78.8	0.0
Townsville	E	39.2	8.0	39.2	63.4	8.0	40.7	100.0	8.0	60.1	63.4	8.0
Cairns	C	36.9	0.0	0.0	50.8	0.0	35.6	100.0	0.0	38.8	50.8	0.0
Mt Isa	A	0.0	0.0	100.0	100.0	0.0	100.0	100.0	0.0	100.0	0.0	0.0

[a] Lower limits on these seven contiguous capacity zones (A to G) are 0, 0.1×10^6, 0.2×10^6, 0.3×10^6, 0.4×10^6, 0.5×10^6 and 1.0×10^6 head a year, respectively.

TABLE 2

Capacities and Capacity Utilisations for the Clean-Slate Models

Region	No seasonal or state of nature variation		Between-year variation only				Seasonal variation only		
	Cap[a]	$T_{\bullet\bullet}$[b]	Cap	$T_{p\bullet}$	$T_{a\bullet}$	$T_{g\bullet}$	Cap	$T_{\bullet p}$	$T_{\bullet o}$
South Queensland	5846	100.0	6833	73.2	100.0	100.0	6073	100.0	83.0
Brisbane	4946	100.0	5000	100.0	100.0	100.0	5000	100.0	88.0
Toowoomba	900	100.0	1833	0.0	100.0	100.0	1073	100.0	59.9
Roma	0	0.0	0	0.0	0.0	0.0	0	0.0	0.0
Maryborough	0	0.0	0	0.0	0.0	0.0	0	0.0	0.0
Central Queensland	2091	100.0	2109	53.0	53.0	100.0	2433	100.0	0.0
Rockhampton	1447	100.0	444	100.0	100.0	100.0	1620	100.0	0.0
Emerald	644	100.0	991	0.0	0.0	100.0	813	100.0	0.0
Mackay	0	0.0	674	100.0	100.0	100.0	0	0.0	0.0
North Queensland	930	100.0	2288	35.0	41.0	100.0	2156	100.0	11.0
Bowen	450	100.0	450	100.0	100.0	100.0	450	100.0	52.9
Townsville	0	0.0	1345	0.0	0.0	100.0	1021	100.0	0.0
Mt Isa	480	100.0	493	72.0	100.0	100.0	685	100.0	0.0
Cairns	0	0.0	0	0.0	0.0	0.0	0	0.0	0.0
Total capacity	8867		11230				10662		

Region	Full stochastic model						
	Cap	T_{pp}	T_{po}	T_{ap}	T_{ao}	T_{gp}	T_{go}
South Queensland	6128	100.0	61.4	100.0	75.7	141.0	96.1
Brisbane	5000	100.0	75.2	100.0	92.7	150.0	100.0
Toowoomba	1128	100.0	0.0	100.0	0.0	102.0	78.8
Roma	0	0.0	0.0	0.0	0.0	0.0	0.0
Maryborough	0	0.0	0.0	0.0	0.0	0.0	0.0
Central Queensland	2538	51.9	18.4	100.0	24.5	109.0	30.3
Rockhampton	900	72.5	0.0	100.0	0.0	100.3	0.0
Emerald	829	0.0	0.0	100.0	0.0	100.0	0.0
Mackay	809	82.1	57.6	100.0	77.0	126.3	95.2
North Queensland	2045	40.5	0.0	100.0	0.0	103.0	0.0
Bowen	450	72.0	0.0	100.0	0.0	115.0	0.0
Townsville	1088	0.0	0.0	100.0	0.0	100.0	0.0
Mt Isa	507	100.0	0.0	100.0	0.0	100.0	0.0
Cairns	0	0.0	0.0	0.0	0.0	0.0	0.0
Total capacity	10771						

[a] Capacity measured as head per day.
[b] Throughput expressed as a percentage.
First throughput subscript: p — poor year; a — average year; g — good year. Second throughput subscript: p — peak season; o — off-peak season.

Allowing only for seasonality leads to capacities more in line with those of the full stochastic model. Capacities were aligned with throughputs in the peak season of the expected year, not unlike the full stochastic model in which capacities were aligned with throughputs in the peak season of the 'average' type of year. As with other models, however, capacity utilisations differ considerably from the full stochastic model. In the peak season, throughputs are similar only to those in the average year of the full stochastic run, while throughputs in the off-peak season are markedly different from those in every type of year in the full stochastic run.

The non-allowance for seasonally high and low supplies in the peak and off-peak seasons resulted in anomalous capacity utilisations and high infeasibility costs at various plants in the long run. Table 3 summarises the normative results of the clean-slate runs. The total costs of each of the solutions based on the assumptions inherent in the respective models are provided. These costs are then compared with the total costs which would be incurred if one used the existing slaughter capacity optimally in the short run, based on the same model assumptions. Thus, for example, the total cost of the deterministic model solution is calculated assuming no seasonal or between-year variability and this is compared with the best deterministic use of the existing facilities. Each comparison represents the estimate of the normative gains that analysis based on the particular assumptions about variability would indicate.[6] Ultimately, however, the actual gains to be made are of interest. A more useful comparison, then, is that of the costs of the alternative location–size patterns (including the existing one) evaluated in the presence of both seasonal and between-year variation. Thus the capacities determined by each of the three simplified models and the existing pattern were fixed and the short-run full stochastic solution obtained for each.

As shown in Table 3, the traditional deterministic plant location model would suggest potential normative gains of around $11m ($216.5m less $205.5m), or a saving of about 5 per cent of total costs. However, implementing this plan would actually result in a cost increase of $1.4m ($217.5m less $218.9m) over the existing situation. The fact that both models allow for only one type of variation also suggests significant cost savings over the existing solution, of the order of $7m or 3 per cent of total costs. Again, as with the deterministic model, the normative gains of both models have been significantly overestimated since the real gains are only about $4m. Nevertheless, the gains from implementing the solution from a model recognising only one source of variation are only $0.5m less than the estimated gains from implementing the solution from the full stochastic model ($4.5m). This is because both models, like the full stochastic model, have to account for extreme throughputs leading to similar optimal capacities in all three models.

Previously, clean-slate plant location models have indicated sizeable normative gains. For example, Cassidy, McCarthy and Toft (1970) estimated gains of 10 per cent of total costs from rationalisation of the

[6] This is strictly not true since one should compare the cost of actual use of existing facilities with costs of actual use of envisaged facilities. We use the costs of the least-cost short-run solution as an estimate of the actual costs.

TABLE 3

Estimated Costs of the Clean-Slate Solutions ($ m)

Estimated costs	No seasonal or state of nature variation (CS1)	Between-year variation only (CS2)	Seasonal variation only (CS3)	Full stochastic (CS4)	Existing situation
Assessed on model assumptions[a]	205.5	209.9	209.4	213.1	216.5
Assessed on assumptions of full stochastic model[b]	218.9	213.5	213.4	213.1	217.5

[a] CS1 assumes no variability; CS2 allows for between-year variability; and CS3 allows for seasonal variability. The cost of the existing situation is assessed on the assumptions of CS1. The cost of the existing situation based on the assumptions of CS2 and CS3 was essentially similar.

[b] That is, assuming full seasonal and between-year variability.

slaughtering facilities serving the eastern central Queensland cattle industry. The Industries Assistance Commission (1983) also estimated gains of the order of 11 per cent for the New South Wales meat processing industry. Ferguson and McCarthy (1970) found potential gains of 40 per cent from rationalising Australian wool handling. But the finding in this paper (that the suggested gains from the traditional clean-slate model of 5 per cent are considerably more than the 2 per cent maximum gains from implementing the full stochastic solution) is sufficient to query whether a significant proportion of the gains estimated in deterministic analyses may be due to misspecification of the model rather than to any real cost savings from rationalisation.

Concluding Comments

Developments are probably necessary in order to maximise the model's usefulness in planning the location of plant facilities. For example, supply and demand arguably are not perfectly inelastic in the long run and should be endogenously determined along with the plant facilities. Attention also needs to be given to modelling the imperfectly competitive behaviour engendered by space.

Notwithstanding its limitations, the model is a marked improvement over deterministic plant location models in terms of the quantity and quality of information it yields. For example, commodity flows, plant utilisations, and regional supply and demand location rents (as outlined by Stevens 1961) can be determined for all possible states of nature rather than an average state of nature. Plant throughput variation costs can be determined directly and these costs may be traded off against transportation and processing costs in a multi-objective analysis.

But more importantly, as evidenced in the empirical work, the implications of the analysis can be quite different. Plant throughputs and commodity shipments obtained from average short-run models may bear no resemblance to those of a short-run stochastic model. For long-run plant location problems, variability may also have significant implications for plant locations, sizes, throughputs and commodity flows. Normative gains associated with long-run industry rationalisation will be overestimated if variation in commodity supply and demand is ignored. The only way to discover any of these differences is to model the variability explicitly for stochastic plant location problems.

APPENDIX

The Discrete Stochastic Plant Location Model

The core of the model may be stated as follows:

(A0) Minimise $\Sigma_p \Sigma_t \Sigma_s A_{tps} X^S{}_{tps} + \Sigma_p \Sigma_j \Sigma_s D_{pjs} X^D{}_{pjs} + \Sigma_p \Sigma_h F_{ph} W_{ph} + \Sigma_p \Sigma_k T_{pk} + \Sigma_p \Sigma_s e_{ps} X^A{}_{ps}$

subject to:

(A1) $\Sigma_p X^S{}_{tps} \leq Q^S{}_{ts}$ for all t and s

(A2) $\Sigma_p X^D{}_{pjs} \geq Q^D{}_{js}$ for all j and s

(A3) $X_{ps} - \Sigma_t X^S{}_{tps} = 0$ for all p and s

(A4) $X_{ps} - \Sigma_j X^D{}_{pjs} = 0$ for all p and s

(A5) $Y_p - \Sigma_i S_{pis} = 0$ for all p and s

(A6) $X_{ps} - \Sigma_i R_{pi} S_{pis} = 0$ for all p and s

(A7) $Y_p - \Sigma_h (CAP)_{ph} W_{ph} = 0$ for all p

(A8) $\Sigma_h W_{ph} \leq 1$ for all p

(A9) $-M \sum_{k=2}^{z} G_{pk} - T_{pi} + \Sigma_s \Sigma_i w_s C_{pli} S_{pis} \leq 0$ for all p

(A10) $MG_{pr} - 2M \sum_{\substack{k=2 \\ k \neq r}}^{z} G_{pk} - T_{pr} + \Sigma_s \Sigma_i w_s C_{pri} S_{pis} \leq M$ for all p and $r = 2, \ldots, z$

(A11) $\Sigma_k G_{pk} \leq 1$ for all p

(A12) $B^c{}_{pv} - \sum_{h=g}^{f} W_{ph} \leq 0$ for all v and p;

where v is an index of regions on the capital cost function at which economies of size become more pronounced; g is the initial point in economies of size region v; and f is the initial point in economies of size region $v+1$

(A13) $\sum_{h=g+1}^{f} W_{ph} - B^c{}_{pv} - PB^c{}_{pv+1} \leq 0$ for all v and p;

and where $P=0$ if $v+1$ is greater than the number of regions in which economies of size become more pronounced for plant p, otherwise $P=1$

(A14) $B^v{}_{pu} - \sum_{i=a}^{b} S_{pis} \leq 0$ for all p, u and s;

where u is the index of regions on the variable processing cost function at which economies of scale become more pronounced; a is the initial defining point in economies of scale region u; and b is the initial defining point in economies of scale region $u+1$

(A15) $\sum_{i=a+1}^{b} S_{pis} - B^{v}_{pu} - NB^{v}_{pu+1} \leq 0$ for all u and p;

and where $N=0$ if $u+1$ is greater than the number of regions in which economies of scale become more pronounced for plant p, otherwise $N=1$

(A16) $E_p - \Sigma_s(1/n)X_{ps} = 0$ for all p

(A17) $-E_p + X_{ps} + X^{A}_{ps} \geq 0$ for all p and s

(A18) $X^{S}_{tps}, X^{D}_{pjs}, X_{ps}, Y_p, W_{ph}, T_{pk}, S_{pis}, E_p, X^{A}_{ps} \geq 0$ for all h, i, j, k, p, s, u and v

(A19) $G_{pk}, B^{c}_{pv}, B^{v}_{psu}$ are binary variables for all k, p, s, u and v

where the variables are defined as follows:

X^{S}_{tps} = number of units of raw product shipped from supply region t to processing plant p in state of nature s;
X^{D}_{pjs} = number of units of final product shipped from processing plant p to demand region j in state of nature s;
X_{ps} = number of units of raw product processed in plant p in state of nature s;
Y_p = processing 'capacity' of plant p;
W_{ph} = the weight attached to the hth defining point of the capacity costs for processing plant p;
T_{pk} = the expected variable processing costs incurred by processing plant p when operating in the kth capacity zone;
S_{pis} = a scaling variable for processing plant p in state of nature s for capacity defining point i;
G_{pk} = binary variable for processing plant p for capacity zone $j, j=2, \ldots, z$;
B^{c}_{pv} = binary variable for processing plant p for the vth region of economies of size;
B^{v}_{psu} = binary variable for processing plant p in state of nature s for the uth region of economies of throughput;
E_p = mean throughput for plant p;
X^{A}_{ps} = negative deviation of throughput from mean throughput for plant p in state of nature s;

and where the following are parameters:

$[F_{ph}, (CAP)_{ph}]$ is a point on the capital cost − capacity function;
e_{ps} is a weighting/cost coefficient on plant throughput deviations;
w_s is the probability of state of nature s;
$[C_{pri}, R_{pi}]$ is a point on the variable processing cost − capacity utilisation relationship for the rth capacity zone;

M is a large positive number;
A_{tps} is a unit assembly cost;
D_{pjs} is a unit distribution cost;
Q^S_{ts} is the supply in region t in state of nature s;
Q^D_{js} is the demand in region j in state of nature s;

and where there are:
S supply regions;
D demand regions;
P processing regions;
n states of nature;
z capacity zones;
h defining points for the capital cost curve for a plant; and
i defining points for the variable processing cost curve for a plant.

Comments

Plant capacity, and the (capital) costs of providing it, are specified through equation (A7) and the third term of the objective function. In general, capital costs would be non-linear functions of capacity, but for modelling purposes, the curves are linearised using a series of linear segments and, where necessary to accommodate economies of size associated integer variables. The methods used in the study for linearisation of the cost curve and use of integer variables to accommodate economies of size or throughput follow the approaches used by Wagner (1969) and Gray (1970).

The total variable processing cost curve is modelled for each state of nature, the costs being defined in equations (A9) to (A11). Because these costs are defined in relation to capacity utilisation, a linkage between throughput and capacity is necessary. This linkage is first described for a single state of nature.

Consider the variable processing cost curve for plant p. It can be approximated by a linear segmented curve represented by the usual linearisation construction:

$$COST_p - \Sigma_i C_{pi} X_{ip} = 0 \tag{A20}$$

$$REL\ UTIL_p - \Sigma_i R_{pi} X_{ip} = 0 \tag{A21}$$

$$\Sigma_i X_{ip} = 1 \tag{A22}$$

$$X_{ip} \geq 0 \qquad \text{for all } i$$

where (R_{pi}, C_{pi}) are the co-ordinates of the segment defining points, the X_{ip} activities (between 0 and 1) define a convex combination of these points, and $COST_p$ and $REL\ UTIL_p$ are activities with obvious meanings.

By multiplying the three equations by capacity, and defining a throughput activity ($X_p = REL\ UTIL_p * Y_p$) for the plant, a capacity activity (Y_p), a total variable processing cost activity ($T_p = COST_p * Y_p$) and scaling variables ($S_{pi} = X_{ip} * Y_p$), one obtains the following equations:

(A23) $$T_p - \Sigma_i C_{pi} S_{pi} = 0$$

(A24) $$X_p - \Sigma_i R_{pi} S_{pi} = 0$$

(A25) $$Y_p - \Sigma_i S_{pi} = 0$$

(A26) $$S_{pi} \geq 0 \quad \text{for all } i$$

This formulation converts any throughput and capacity to total variable processing costs by scaling the unit variable processing costs for the particular relative utilisation by capacity.

The linkages modelled in equations (A23) to (A25) relate only to a single state of nature. To allow for multiple states of nature, the throughput and scaling variables must be defined for each state of nature. The equations (A24) and (A25) for scaling the variable processing cost curve must also be repeated. On the other hand, equation (A23) need be merely extended by defining T_p as the expected total variable costs incurred by the plant, and by summing probability-weighted variable processing cost activities (S_{pi}) over all segments and all states of nature.

(A27) $$T_p - \Sigma_s \Sigma_i w_s C_{pi} S_{pis} = 0$$

Finally, to allow for alternative capacity zones, the cost defining equation (A27) must be replaced with a set of cost equations (A9) and (A10). (Note: for each capacity zone (r), both the cost (C_{pri}) and, in theory, the capacity utilisations (R_{pri}) may vary. However, variable processing cost curves for all capacity zones may be based on a single set of R_{pi} values defined as the union of the sets for all capacity zones. This enables the segment defining points to be represented by a single set of activities (X_{pi}), each defining particular capacity utilisations (R_{pi}), with variation arising from different capacity zones being represented by changes in variable processing costs (C_{pri}).) The resulting cost equations are incorporated in the model as rows (A9) and (A10). Each row relates to a particular capacity zone and defines the cost for that zone as in equation (A27), but as well includes binary variables which, along with the convexity constraint (A11), serve to ensure that only the one row corresponding to the relevant capacity zone yields a cost contribution in the objective row.

If the capacity of plant p is determined by the model to be in capacity zone 1, then the variable processing costs associated with this zone (C_{pli}, $i = 1, \ldots, n$) will be the relevant costs. Every G_{pk} will be set to zero and, in seeking to minimise costs, only T_{pl} will be an active (non-zero) transfer cost activity. For capacity zone $k (k > 1)$, G_{pk} will be set to unity and T_{pk} alone will be active. A similar mechanism operates when any other capacity zone is relevant.

References

Anderson, J. R. (1979), 'Impacts of climatic variability in Australian agriculture: a review', *Review of Marketing and Agricultural Economics* 47(3), 147–77.

Balachandran, V. and Jain, S. (1976), 'Optimal facility location under random demand with general cost structure', *Naval Research Logistics Quarterly* 23, 421–36.

Brown, C. G. (1985), *PLOMAG: A General Plant Location Matrix Generator*, Agricultural Economics Discussion Paper 1/85, Department of Agriculture, University of Queensland.

Cassidy, P. A., McCarthy, W. O. and Toft, H. I. (1970), 'An application of spatial analysis to beef slaughter plant location and size, Queensland', *Australian Journal of Agricultural Economics* 14(1), 1–20.

Charnes, A., Cooper, W. W. and Thompson, G. L. (1965), 'Constrained generalized medians and hypermedians as deterministic equivalents for two stage linear programs under uncertainty', *Management Science* 12(1), 83–112.

Cocks, K. D. (1968), 'Discrete stochastic programming', *Management Science* 15(1), 72–9.

Control Data Corporation (1979), *APEX III Reference Manual Version 1.2*, Data Services Publication, Minneapolis.

Ferguson, D. C. and McCarthy, W. O. (1970), 'A spatial analysis of size and location of Australian wool selling centres', *Review of Marketing and Agricultural Economics* 38(4), 153–69.

Gray, P. (1970), 'Exact solution of the site selection problem by mixed integer programming', in E. M. L. Beale (ed.), *Applications of Mathematical Programming Techniques*, English University Press, London, 261–70.

Hazell, P. B. R. (1971), 'A linear alternative to quadratic and semi variance programming for farm planning under uncertainty', *American Journal of Agricultural Economics* 53(1), 53–62.

Hurt, V. G. and Tramel, T. E. (1965), 'Alternative formulations of the transhipment problem', *Journal of Farm Economics* 47(3), 763–73.

Industries Assistance Commission (1983), *Report into the Meat Processing Industry and Abattoir Sector*, AGPS, Canberra.

Kilmer, R. L., Spreen, T. and Tilley, D. S. (1983), 'A dynamic plant location model: the east Florida fresh citrus packing industry', *American Journal of Agricultural Economics* 65(4), 730–7.

King, G. A. and Logan, S. H. (1964), 'Optimal location, number and size of processing plants with raw product shipments', *Journal of Farm Economics* 46(1), 94–108.

de Leeuw, F. (1962), 'The concept of capacity', *Journal of the American Statisatical Association* 57(300), 826–40.

Ryland, G. J. and Guise, J. W. B. (1975), 'A spatiotemporal quality competition model of the Australian sugarcane processing industry', *American Journal of Agricultural Economics* 57(3), 431–8.

Stevens, B. H. (1961), 'Linear programming and location rent', *Journal of Regional Science* 3(2), 15–26.

Wagner, H. M. (1969), *Principles of Operations Research*, Prentice Hall, Englewood Cliffs.