"The Methods Of Statistical Mechanics And Its Applications To Environmental Problems"

SSERTATION PRESENTED IN PART-REQUIREMENT
R THE DEGREE OF MASTERS OF SCIENCE IN
THEMATICS AT THE UNIVERSITY OF ZAMBIA.

rick A. Phiri
versity of Zambia

tember, 1976.

ACKNOWLEDGEMENTS

I would like to thank the Mathematics Department for initiating the M.Sc. programme and the University of Zambia for financing it. I am also very grateful to Dr. R. A. Ampomah and Dr. S. I. Emenalo for their helpful advice and wise counsel. Mention should also be made of Professor E. F. Bartholomeusz, Dr. A. Ciampi and Dr. N. Murthy for helping to make the M.Sc. programme a success.

SECTION 1 : INTRODUCTION

In recent years, mathematical models have been constructed in the study of environmental problems. These models have been achieved on the basis of analogies with statistical mechanical techniques. The methods used in both urban studies and statistical mechanics have been unified by the concepts of information theory. The underlying unification is analysed by Wilson [8].

Models, derived from the maximisation of an entropy function, can be constructed to study, say, a transport system, residential location, and commodity flows in a specified region. Similar models can be built in a linear programming framework, and it has been shown by Wilson and Senior [12], that the linear programming model is the limit of the entropy maximising model as certain parameters tend to infinity. Wilson and Senior [5] proved this hypothesis true for the residential location case using data for the City of Leeds.

In Section 2 of this dissertation the canonical ensemble method for analysing systems of particles in statistical mechanics is illustrated. Information Theory and its role in statistical mechanics and mathematical modelling of urban systems is outlined in Section 3. A case study made of urban transport flows following, in the main, the work of Wilson [9] is discussed in Section 4. Section 5 is a review of the other areas in urban studies where the methods of statistical mechanics may be applied. Some concluding observations are made in Section 6.

2. THE CANONICAL ENSEMBLE METHOD

Consider a thermodynamic assembly consisting of N localized and independent systems. Specify the states of the assembly by n_i and let each state have permissible energies e_i. Then

$$\sum_i n_i = N , \qquad (1)$$

and

$$\sum_i n_i e_i = E , \quad i = 1, 2, 3, \ldots, \qquad (2)$$

where E is the total energy of the N systems.

Suppose we are to predict the number of systems that have specified energies on the basis of the information contained in equations (1) and (2). This task would not be possible unless we had knowledge of the precise positions and velocities of the systems under consideration. We can, however, following Tolman [7], deduce from the given data what the most probable distribution of the energies of the different states of the assembly are.

We start by evaluating Ω, the number of distinguishable states of the assembly, assuming that all the states of the assembly are equally probable. The number of different states of the assembly corresponding to a given set of numbers $\{n_i\}$ is t, where

$$t = \frac{N!}{\prod_i n_i!} , \qquad (3)$$

so that

$$\Omega = \sum_{\substack{\text{(all possible}\\ \text{sets of values}\\ \{n_i\})}} \frac{N!}{\prod_i n_i!} . \qquad (4)$$

Since, when N is large ($\sim 10^{23}$), as is the case in thermodynamic assemblies, only the largest term in (4) makes an effective contribution to Ω, it is sufficient to obtain the greatest value of t for any permissible values of $\{n_i\}$. This is achieved by finding those n_i's that maximise t subject to (1) and (2).

We thus form the Lagrangian Z, such that

$$Z = \ln t + \alpha(\sum_i n_i - N) + \beta(\sum_i n_i e_i - E), \qquad (5)$$

where α and β are the Lagrange multipliers associated with equations (1) and (2) respectively. In equation (5), it is more convenient to maximise $\ln t$ rather than t, and then it is possible to use Stirling's approximation

$$\ln N! = N \ln N - N. \qquad (6)$$

In fact, in maximising $\ln t$ rather than t, no generality is lost since $\ln t$, being a monotonic function, attains its maximum at the point at which t is maximum.

The n_i's which maximise Z are the solutions of

$$\frac{\partial Z}{\partial n_j} = 0, \qquad (7)$$

Note that

$$\frac{\partial}{\partial n_i} t = \frac{\partial}{\partial n_j}(\ln N! - \sum_i \ln n_i!),$$

$$= - \sum_i \ln n_i,$$

using Stirling's approximation (6). Equation (7) may then be written as

$$\frac{\partial Z}{\partial n_i} = - \ln n_i + \alpha + \beta e_i = 0,$$

$$n_i = e^{\alpha + \beta e_i} \quad . \tag{8}$$

Equations (8) give the most probable distribution of the numbers n_i satisfying (1) and (2). That is the assembly is more likely to be found in a state corresponding to the distribution numbers given in (8) than in any other state, given that the states of the assembly are equally probable.

In the above illustration, a deductive prediction, given by (8), has been made about a system whose data $\{(1), (2)\}$ is incomplete. In other words, with the given information only, the most reasonable predictions have been made as to what states of the assembly are most likely to occur. Such statistical mechanical predictions contain an element of uncertainty, and are based on inductive reasoning. Since the formalism for inductive reasoning is probability theory, it follows that probability theory plays a major role in statistical mechanics. This is manifest by the fact that, given a system, statistical mechanics gives the probability distribution of the most likely outcome of occurrences in the specified systems.

3. INFORMATION THEORY

We noted above that statistical mechanical probability distributions contain an element of uncertainty. Shannon and Weaver [6] propounded a theory, called Information Theory, that gives the quantitative measure of the missing information in a probability distribution. Below, the Theory is outlined and its role in statistical mechanics investigated. Later, the use of Information Theory in the formulation of mathematical models for the solutions of environmental probalems is illustrated.

In finding an expression for the quantitative measure of the missing information in a probability distribution, we begin by considering the properties a measure of information content of a probability assignment should satisfy, and then derive the unique expression for it. This is achieved by following the Uniqueness Theorem approach. The Uniqueness Theorem gives the expression of the information content of a probability assignmemt given the prior probabilities. The derived expression of information content is then used to find an expression for the quantitative measure of the missing information in a probability distribution. Hobson's approach [4] will be followed.

Information

It is desirable at this point to quantify what is meant by "information". This is done by giving two examples.

Example 1:-

Consider a die with six sides. Let Case (a) be the reduction in the number of equally likely sides from 6 to 3, and Case (b) be the reduction in the number of equally likely sides from 6 to 5. Then we say that the reduction in Case (a) contains more information than the reduction in Case (b).

Example 2:-

Suppose that we want to predict the number of spots showing on the next trial in a die-throwing experiment, given only that the die has six sides with i spots on the i^{th} side, and that the average number of spots obtained in a previous long series of throws was 3.5, say. That is, for the events f_i, we seek a probability assignment $P(\{f_i\}) = p_i$, $(i = 1, 2, \ldots 6)$. Since the p_i are probabilities,

$$\sum_i p_i = 1 , \qquad (9)$$

and

$$\sum_i i\, p_i = 3.5 , \qquad (10)$$

since the average throw is 3.5 .

A probability assignment that fits (9) and (10) is

$$p_4 = \frac{1}{2}, \quad p_3 = \frac{1}{2}, \quad p_1 = p_2 = p_5 = p_6 = 0 \qquad (11)$$

but this assignment seems to assume arbitrarily that f_1, f_2, f_5, f_6, cannot occur, whereas the data does not imply this. Thus, (11) contains more information than is actually given by the data.

Information Content

Let $$I(P; P^o) = I(p_1, p_2, \ldots, p_n ; p_1^o, p_2^o, \ldots, p_n^o) \qquad (12)$$

represent the information content in the probability assignment P relative to the initial probability assignment P^o , where

$$P(\{\xi_i\}) = p_i \ , \qquad (13)$$

and
$$P^o(\{\xi_i\}) = p_i^o \ , \qquad i = 1, 2, \ldots, n \qquad (14)$$

for the set of events ξ_i on a finite sample space S. In seeking a function of the form (12), defined for the probability assignments P and P^o on the finite sample space S, it is reasonable to postulate the properties such a function should satisfy, and then derive the desired function.

<u>Properties of $I(P; P^0)$</u>

(a)
$$I(p_1, p_2, \ldots p_n ; \quad p_1^0 , p_2^0 , \ldots, p_n^0) \qquad (15)$$

is a continuous function; that is, small changes in P and P^0 do not appreciably change the information.

(b)
$$I(p_1, \ldots, p_j, \ldots, p_k, \ldots, p_n ; \quad p_1^0, \ldots, p_j^0, \ldots, p_k^0, \ldots p_n^0)$$

$$= I(p_1, \ldots, p_k, \ldots, p_j, \ldots, p_n ; \quad p_1^0, \ldots, p_k^0, \ldots, p_j^0, \ldots p_n^0) \ , \qquad (16)$$

(c)
$$I(P; P) = 0 \ , \qquad (17)$$

which means that no information is obtained if the final and prior probabilities are the same

(d) for any integers n and n_o for which $n_o \geq n$

$$I(\frac{1}{n}, \ldots, \frac{1}{n}, \ 0, \ldots, 0 ; \quad \frac{1}{n_o}, \ldots, \frac{1}{n_o})$$

is an increasing function of n_o and a decreasing function of n.

(e) $I(p_1, \ldots, p_r, p_{r+1}, \ldots, p_n ; \quad p_1^0, \ldots, p_r^0, p_{r+1}^0, \ldots, p_n^0)$

$$= I(q_1, q_2; \; q_1^0, q_2^0) + q_1 \, I(\frac{p_1}{q_1}, \ldots, \frac{p_r}{q_1} ; \frac{p_1^0}{q_1^0}, \ldots, \frac{p_r^0}{q_1^0})$$

$$+ q_2 \, I(\frac{p_{r+1}}{q}, \ldots, \frac{p_n}{q_2} ; \frac{p_{r+1}}{q_2^0}, \ldots, \frac{p_n^0}{q_2^0}) \qquad (19)$$

Postulate (16) implies that the manner in which the outcomes are labelled does not affect the information, and postulate (19) is known as the composition rule.

Divide the sample space S into two disjoint subspaces S_1 and S_2. Then S_1 and S_2 are events having final probabilities, say,

$$q_1 = p_1 + \ldots + p_r \quad \text{and} \quad q_2 = p_{r+1} + \ldots + p_r$$

respectively, and prior probabilities

$$q_1^0 = q_1^0 + \ldots + p_r^0 \quad \text{and} \quad q_2^0 = p_{r+1}^0 + \ldots + p_n^0 .$$

From Probability theory, it is known that the conditional probability of an event B given the probability of an event A, $P(B|A)$, is given by

$$P(B|A) = \frac{P(B \cap A)}{P(A)} , \qquad P(A) > 0 ,$$

so that the final probabilities of ξ_i, given that $\xi_i \in S_1$ are

$$P(\{\xi_i\}|S) = \frac{P(\{\xi_i\} \cap S)}{P(S)} = \frac{P(\{\xi_i\})}{P(S)} = \frac{p_i}{q_1} .$$

Also, the prior probabilities of ξ_i , given that $\xi_i \in S_1$ are

$$P^0(\{\xi_i\}|S) = \frac{p_i^0}{q_1^0} .$$

Similarly, the final and prior probabilities of ξ_i given S_2 are

$$P(\{\xi_i\}|S_2) = \frac{p_i}{q_2} ,$$

and
$$P^0(\{\xi_i\}|S_2) = \frac{p_i^0}{q_2^0} ,$$

respectively.

Information about the outcome may be given either by specifying the probabilities $p_1, p_2, \ldots p_n$ directly, that is, diagramatically,

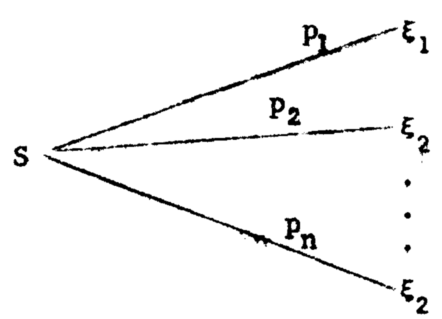

Figure 1. Diagramatic representation of the probability assignment $p_1, p_2, \ldots p_n$.

or by specifying the probabilities q_1, q_2 of the sub-spaces S_1, S_2 and then giving the conditional probabilities $\frac{p_i}{q_1}, \frac{p_i}{q_2}$ within these sub-spaces, which may be diagramatically represented as

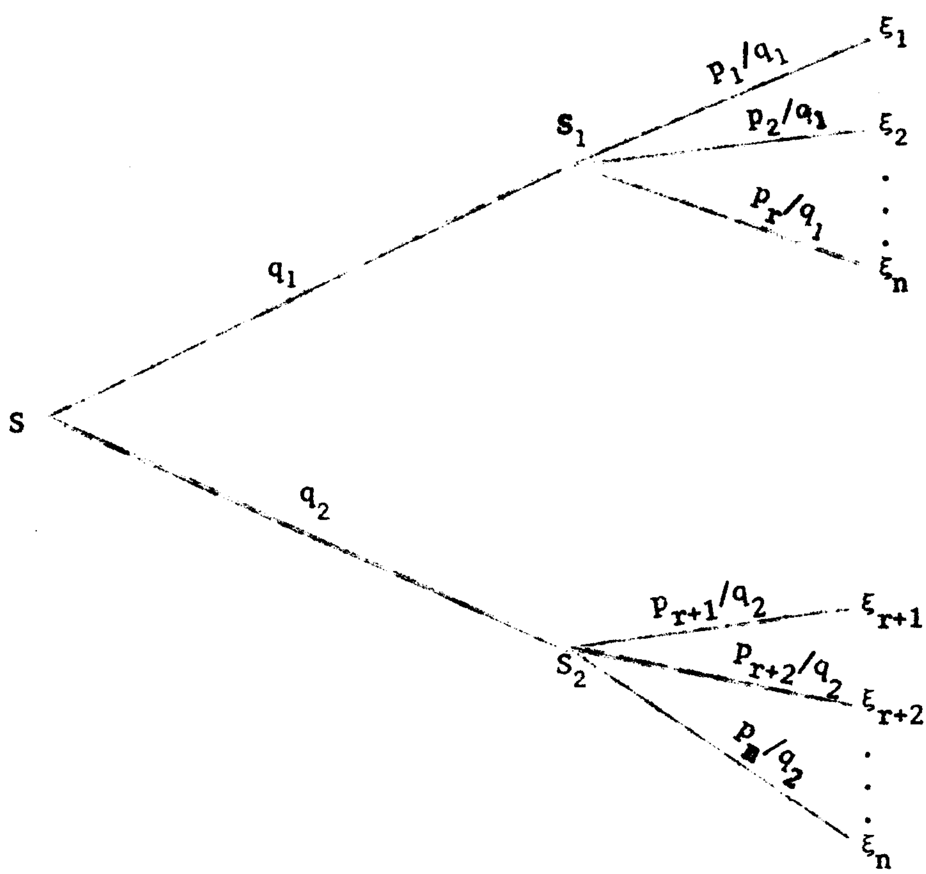

Figure 2. An alternative representation of the probability assignment shown in Figure 1.

Equation (19) asserts that the amount of information represented in the schemes of Figures 1 and 2 are equal, and that the right hand side of (19) is a reasonable expression for the information in Figure 2.

The Uniqueness Theorem

Let $I(p_1, p_2, \ldots p_n ; \; p_1^0, p_2^0, \ldots, p_n^0)$ be a function defined for any pair of probability assignments P, P^0 on a finite sample space $S = \{\xi_1, \xi_2, \ldots, \xi_n\}$ for any n. If this function satisfies postulates (a), (b), (c), (d) and (e), then

$$I(p_1, p_2, \ldots, p_n ; \; p_1^0, p_2^0, \ldots, p_n^0) = \Sigma \, p_i \ln \frac{p_i}{p_i^0} . \tag{20}$$

The reader is referred to Hobson [4] for a proof of the Uniqueness Theorem.

Uncertainty or Missing Information

The expression for the information content given by (20) will be used to derive an expression for the missing information in a probability assignment. It is the expression that we shall derive that will be made use of in our study of mathematical models of environmental studies.

Let the probability assignment corresponding to the maximum knowledge which can be obtained about the outcome of an event be denoted by P^m. Then $I(P^m;P^o)$ is the maximum information obtainable relative to the prior probabilities P^o. If one's actual state of knowledge is described by the probability assignment P, then the missing information needed to attain the state P^m is

$$S = I(P^m ; P^o) - I(P ; P^o) . \tag{21}$$

Therefore $$S = \sum_i p_i^m \ln \frac{p_i^m}{p_i^o} - \sum_i p_i \ln \frac{p_i}{p_i^o} , \tag{22}$$

using (20). In (22),

$$P^m(\{\xi_i\}) = p_i^m ,$$

$$P(\{\xi_i\}) = p_i ,$$

and $$P^o(\{\xi_i\}) = p_i^o .$$

We shall apply the concepts of information theory to situations where the sample space is a finite set, say, $\xi_1, \xi_2, \ldots, \xi_n$, and P^o and P^m are given by

$$P^o(\{\xi_i\}) = \frac{1}{n}, \quad (i = 1, 2, \ldots, n), \tag{23}$$

$$P^m(\{\xi_i\}) = \delta_{ik}, \quad (i = 1, 2, \ldots, n \,;\, k \text{ fixed}). \tag{24}$$

Then (22) with equations (23) and (24) imply

$$S = -\sum_i p_i \ln p_i \,. \tag{25}$$

Equation (25) is our expression for the missing information or uncertainty in the probability assignment P.

Jaynes' Principle for countable sample spaces

"Let $S = \{\xi_1 \ \xi_2, \ldots\}$ be a set of possible outcomes in some experiment, and assume that the prior description of the experiment is symmetric with respect to the ξ_i. If data D is then given concerning the experiment, the probability assignment $P = (p_1, p_2, \ldots)$ which represents D must maximise the uncertainty (25) with respect to all P satisfying D." [4]

What Jaynes' Principle implies is that the unique probability assignment P that contains the given data and does not arbitrarily assume anything more than the given data, is that assignment which maximises the uncertainty function (25), provided that the prior probabilities are symmetrical. Our prior probabilities, given by equation (23), are symmetrical, and so Jaynes' Principle may be applied.

4. TRIP DISTRIBUTION : An Application of Information Theory

The term ln t in equation (5) is normally referred to as entropy in statistical mechanics [6] whereas the information theorist's definition of entropy is given by (25). In view of the discussion in the previous section, we derive an expression for the most probable trip distribution (see below) in a specified region.

Trip Distribution Defined

Suppose a region is divided into N mutually exclusive zones called origins. For a single trip purpose, say shopping, O_i is the total number of trips observed to emanate from origin i during some finite period of time. Let there be M mutually exclusive places called destinations, and let D_j be the total number of trip destinations at j. Denote by T_{ij}, the number of trips from origin i to destination j, c_{ij}, the total value of all resources used in the trip appearing in the traveller's accounts, and T the total number of trips. The quantities of interest are tabulated below.

Destinations

	$i \backslash j$	1	2	. . .	M	$\sum_{j=1}^{M} T_{ij}$
O	1	T_{11}	T_{12}	. . .	T_{1M}	O_1
r	2	T_{21}	T_{22}	. . .	T_{2M}	O_2
i	.	.	.		.	.
g	.	.	.		.	.
i	.	.	.		.	.
n s	N	T_{N1}	T_{N2}	. . .	T_{NM}	O_N
	$\sum_{j=1}^{N} T_{ij}$	D_1	D_2	. . .	D_M	$T = \sum_{j=1}^{N} O_i = \sum_{j=1}^{M} D_j$

Figure 3. Trip Distribution Table.

The key assumption in our derivation of the most probable trip distribution is that all states of the system, as depicted in Figure 3, are equally probable. It is then possible to apply Jaynes' Principle which was stated in Section 3.

Define p_{ij} such that

$$p_{ij} = \frac{T_{ij}}{T} \quad . \tag{26}$$

Then, equation (25) becomes

$$S = -\sum_i \sum_j p_{ij} \ln p_{ij} \quad . \tag{27}$$

The constraint equations to be satisfied are, from Figure 3,

$$\sum_j T_{ij} = O_i , \tag{28}$$

and

$$\sum_i T_{ij} = D_j . \tag{29}$$

A third constraint has to be satisfied:

$$\sum_i \sum_j T_{ij} c_{ij} = C , \tag{30}$$

where C is the total expenditure in the system.

Using (26), equations (28), (29) and (30) may be written in the form

$$\sum_j p_{ij} = \frac{O_i}{T} , \tag{31}$$

$$\sum_i p_{ij} = \frac{D_j}{T} , \tag{32}$$

$$\sum_i \sum_j p_{ij} c_{ij} = \frac{C}{T} . \tag{33}$$

Following Jaynes' Principle, we proceed by maximising (27) subject to (31), (32) and (33). We form the Lagrangian, L, in the usual way,

$$L = - \sum_i \sum_j p_{ij} \ln p_{ij} + \lambda_i \left(\frac{O_i}{T} - \sum_j p_{ij}\right)$$

$$+ \lambda_j^{(2)} \left(\frac{D_j}{T} - \sum_i p_{ij}\right) + \beta\left(\frac{C}{T} - \sum_i \sum_j p_{ij} c_{ij}\right) , \qquad (34)$$

where λ_i , $\lambda_j^{(2)}$ and β are Lagrange multipliers associated with the constraint equations (31), (32) and (33). The p_{ij}'s which maximise L, and which therefore give rise to the most probable distribution of trips, are the solutions of

$$\frac{\partial L}{\partial p_{ij}} = 0 , \qquad (35)$$

and the constraint equations (31), (32) and (33).

Then

$$\frac{\partial L}{\partial p_{ij}} = - \ln p_{ij} - 1 - (\lambda_i + \lambda_j^{(2)} + \beta c_{ij}) = 0 \qquad (36)$$

so that

$$p_{ij} = e^{-1 - (\lambda_i + \lambda_j^{(2)} + \beta c_{ij})} .$$

Put $1 + \lambda_i = \lambda_i^{(1)}$ to obtain

$$p_{ij} = e^{-(\lambda_i^{(1)} + \lambda_j^{(2)} + c_{ij})} \, . \tag{37}$$

Substitute (37) in (31) and (32):

$$\sum_j p_{ij} = e^{-\lambda_i^{(1)}} \sum_i e^{-(\lambda_j^{(2)} + \beta c_{ij})} = \frac{O_i}{T} \, ,$$

from which

$$e^{-\lambda_i^{(1)}} = \frac{O_i}{T \sum_j e^{-(\lambda_j^{(2)} + \beta c_{ij})}} = \frac{O_i \, A_i}{T} \, , \text{ say.} \tag{38}$$

$$\sum_i p_{ij} = e^{-\lambda_j^{(2)}} \sum_i e^{-(\lambda_i^{(1)} + \beta c_{ij})} = \frac{D_j}{T} \, ,$$

so that

$$e^{-\lambda_j^{(2)}} = \frac{D_j}{T \sum_i e^{-(\lambda_i^{(1)} + \beta c_{ij})}} = D_j \, B_j^* \, , \text{ say.} \tag{39}$$

where, by equations (38) and (39),

$$A_i = \frac{1}{\sum_j e^{-({}_j^{(2)} + c_{ij})}} , \tag{40}$$

and
$$B_j^* = \frac{1}{T \sum_i e^{-(\lambda_i^{(1)} + \beta c_{ij})}} . \tag{41}$$

Substituting (38) and (39) in (37) gives

$$p_{ij} = A_i B_j^* \frac{O_i}{T} D_j e^{-\beta c_{ij}} , \tag{42}$$

which is the required expression for the maximum probability distribution of p_{ij} subject to the constraints (31), (32) and (33). To obtain an estimate of the number of trips between zones i and j substitute (26) in (42) to obtain

$$T_{ij} = A_i B_j O_i D_j e^{-\beta c_{ij}} . \tag{43}$$

Equation (43) may be obtained by maximising $\ln t$, where

$$t = \frac{T!}{\prod_i \prod_j T_{ij}!} \tag{44}$$

subject to (28), (29) and (30). This is the procedure often followed by Wilson [9], [11], and Wilson and Senior [5], [12]. The procedure followed in deriving (42), called the entropy maximising procedure, is theoretically more preferable to the method often adopted by Wilson and Senior, called the probability maximising procedure, because it is independent of the choice of units for T_{ij} . The procedure works with p_{ij} which, as given by (26), is dimensionless. The advantage of the entropy maximising procedure over the probability maximising procedure is better appreciated in the study of inter-regional commodity flows, and in the study of energy flows in an urban space economy.

Gravity Models

It has been stated above that Wilson et al used the probability maximising technique to obtain trip distribution models. For the specific case outlined above, the trip distribution model is given by (43), where

$$A_i = \frac{1}{\sum_j B_j D_j e^{-\beta c_{ij}}} , \tag{45}$$

and

$$B_j = \frac{1}{\sum_i A_i O_i e^{-\beta c_{ij}}} . \tag{46}$$

The details of the derivation can be found in Wilson [9].

It would be convenient to introduce the gravity model at this point. Gravity models are a family of four sub-models concerned with trip generation, distribution, model split and assignment, where each sub-model is appropriate to different circumstances. It follows that the model given by equations (43), (45) and (46) may be referred to as the trip-distribution gravity model.

In order to obtain a better understanding of the trip-distribution model, the parameters that appear in the model are explained.

Balancing Factors

A_i and B_j in the trip-distribution model are called balancing factors. This is because they modify the trip matrix (Figure 3) so that (28) and (29) are satisfied. They are obtained by solving equations (45) and (46) iteratively. One would be interested to know whether the iterative procedure converges to a unique solution. There are different methods of iteration. Evans [3] has shown that the methods converge, but uniqueness of the solutions varies with each method.

The interpretation of A_i and B_j is explained next. Referring to O_i and D_j as productions and attractions respectively, the model represented by (43), (45) and (46) may be called the production-attraction constrained model. This is because T_{ij} is constrained to add up to an independently given number of productions and attractions.

Suppose O_i's are not independently given so that only (29) has to be satisfied. Then the model can take the form

$$T_{ij} = B_j W_i^{(1)} D_j e^{-\beta c_{ij}} , \tag{47}$$

where
$$B_j = \frac{1}{\sum_i W_i^{(1)} e^{-\beta c_{ij}}} , \tag{48}$$

and $W_i^{(1)}$ represents an "emissiveness" index of zone i. Substituting (48) in (47) gives

$$T_{ij} = D_j \frac{W_i^{(1)} e^{-\beta c_{ij}}}{\sum_i W_i^{(1)} e^{-\beta c_{ij}}} . \tag{49}$$

The term $W_i^{(1)} e^{-\beta c_{ij}}$ can be interpreted as the emissivity of residents in zone i for shops in zone j. Thus $\frac{1}{B_j}$ or $\sum_i W_i^{(1)} e^{-\beta c_{ij}}$ represents the total emissivity of residents in i for shops in j.

If the D_j's are not independently given, then only (28) need be satisfied and the model can be written as

$$T_{ij} = A_i O_i W_j^{(2)} e^{- c_{ij}} , \tag{50}$$

where
$$A_i = \frac{1}{\sum_j W_j^{(2)} e^{-\beta c_{ij}}} , \tag{51}$$

and $W_j^{(2)}$ is an "attractiveness" index. On the same lines as for the interpretation of B_j, $\frac{1}{A_i}$ or $\sum_j W_j^{(2)} e^{-\beta c_{ij}}$ represents the total accessibility for residents of zone i to shops in zone j. Furthermore, it can be easily seen that as c_{ij} increases, the accessibility decreases, and as c_{ij} decreases, the accessibility increases.

From the discussion above, it is important to note that it is not possible to assign values of the A_i 's and B_j's independently of each other. This is because the emissibility effects operate simultaneously.

The Impedance Function

The cost decay function $e^{-\beta c_{ij}}$ that appears in equations (42) and (43) is called the impedance function. This is because the travel costs, c_{ij}, may be considered as a travel impedance or deterrence.

Up to now this has arisen quite naturally as a negative exponential function. This follows directly from our assumption (30) or (33), that the total expenditure in the system is constrained to a fixed quantity. However, this is true only for a given fixed transport system, for if the transport infrastructure changed, β would have to be adjusted.

Since, generally, the impendance function will decrease as c_{ij} increases, one would consider situations where the function $c_{ij}^{-\beta}$ gives a better fit than $e^{-\beta c_{ij}}$. Then, in our derivation, c_{ij} would be replaced by $\ln c_{ij}$ so that $e^{-c_{ij}}$ is replaced by

$$e^{-\beta \ln c_{ij}} = c_{ij}^{-\beta} . \tag{52}$$

This means that while c_{ij} will remain the measured costs, travellers actually perceive the costs in a manner varying like $\ln c_{ij}$. For urban studies, where the trip costs are generally small $e^{-\beta c_{ij}}$ would give a better fit, whereas for regional studies $c_{ij}^{-\beta}$ would be used.

Trip-end Estimates

The O_i's and D_j's are the trip-end estimates. They are normally obtained from a separate model, a trip-generation model. For the purposes of the trip-distribution model, we will assume these to have been independently given.

5. OTHER AREAS OF URBAN STUDIES

In Section 4 the trip-distribution model was described in some detail. This section will only review the other areas of urban studies where the entropy-maximising technique and the probability-maximising procedure have found applications.

The first application to be discussed will be the disaggregated spatial-interaction model of residential location. Then we will look at models of inter-regional commodity flows and discuss energy and material flows in an urban space economy.

Disaggregated Spatial-Interaction Model for Residential Location

A disaggregated residential location model can be obtained by choosing T_{ij}^{kw} to maximise

$$S = - \sum_{ijkw} \ln T_{ij}^{kw} ! , \qquad (53)$$

subject to

$$\sum_{jw} T_{ij}^{kw} = H_i^k , \qquad (54)$$

$$\sum_{ik} T_{ij}^{kw} = E_j^w , \qquad (55)$$

and

$$\sum_{ijkw} T_{ij}^{kw} (b_{ij}^{kw} - c_{ij}^{w}) = Z , \qquad (56)$$

where Z takes a suboptimal value [5] . The variables can be defined as follows:

T_{ij}^{kw} = the number of type w households living in type k houses in zone i , working in wage jobs in zone j .

H_i^k = the stock of type k houses available in zone i .

b_{ij}^{kw} = the bid price of a type w household (in job location j) for a type k house in zone i ,

c_{ij}^{w} = the average proportion of income which a w-income household spends on housing.

We form the Lagrangian L in the form

$$L = -\sum_{ijkw} \ln T_{ij}^{kw}! + \lambda_i^{k(1)} \left(\sum_{ijkw} T_{ij}^{kw} - H_i^k\right) + \lambda_j^{w(2)} \left(E_j^w - \sum_{ijkw} T_{ij}^{kw}\right)$$

$$+ \mu \left[Z - \sum_{ijkw} T_{ij}^{kw} (b_{ij}^{kw} - c_{ij}^{w})\right] \quad (57)$$

Using Stirling's approximation (6), we obtain

$$\frac{\partial L}{\partial T_{ij}^{kw}} = -\ln T_{ij}^{kw} + \lambda_i^{k(1)} - \lambda_j^{w(2)} + \mu(b_{ij}^{kw} - c_{ij}^{w}) .$$

$\lambda_i^{k(1)}$, $\lambda_j^{w(2)}$ and μ are Lagrange multipliers associated with (54), (55) and (56) respectively.

$\frac{\partial L}{\partial T_{ij}^{kw}} = 0$, gives

$$T_{ij}^{kw} = e^{-\lambda_i^{k(1)}} \; e^{\lambda_j^{w(2)}} \; e^{\mu(b_{ij}^{kw} - c_{ij}^{w})} . \tag{58}$$

Let $$e^{-\lambda_i^{k(1)}} = A_i^k \, H_i^k , \tag{59}$$

and $$e^{\lambda_j^{w(2)}} = B_j^w \, E_j^w . \tag{60}$$

Then (58) can be written as

$$T_{ij}^{kw} = A_i^k \, B_j^w \, H_i^k \, E_j^w \; e^{\mu(b_{ij}^{kw} - c_{ij}^{w})} . \tag{61}$$

Equations (54) and (55) with (61) give

$$A_i^k = \frac{1}{\sum_{jw} B_j^w \, E_j^w \; e^{\mu(b_{ij}^{kw} - c_{ij}^{w})}} \tag{62}$$

and

$$B_j^w = \frac{1}{\sum_{ik} A_i^k \, H_i^k \; e^{\mu(b_{ij}^{kw} - c_{ij}^{w})}} \tag{63}$$

Equations (61), (62) and (63) closely resemble equations (43), (45) and (46) of the trip-distribution model and can be solved iteratively.

The entropy maximising model outlined above may be viewed as a suboptimal version of the following linear programming model:

Choose T_{ij}^{kw} to maximise

$$Z = \sum_{ijkw} T_{ij}^{kw} (b_{ij}^{kw} - c_{ij}^{w}) \quad , \tag{64}$$

subject to

$$\sum_{jw} T_{ij}^{kw} = H_I^{k} \quad , \tag{65}$$

and

$$\sum_{jw} T_{ij}^{kw} = E_j^{w} \quad . \tag{66}$$

The dual problem can be formulated as follows:

Choose α_i^k and ν_j^w to minimise

$$Z' = \sum_{ik} \alpha_i^k H_i^k - \sum_{jw} \nu_j^w E_j^w \quad , \tag{67}$$

such that

$$\alpha_i^k - \nu_j^w \geq b_{ij}^{kw} - c_{ij}^{w} \quad . \tag{68}$$

α_i^k and ν_j^w are the dual variables.

The method of solution is now briefly outlined. The parameter μ in the entropy-maximising model {equations (61), (62) and (63)} is calibrated using bid-rent levels. This means that using data from a previous survey, μ is varied until the predicted level of bid-rents $z^{(p)}$ is a sufficiently close approximation to the observed bid-rents $z^{(obs)}$ for that particular survey. The model can then be used for prediction assuming that the calculated value of μ does not vary with time.

Senior and Wilson [5] ran the models using data for the City of Leeds. The entropy maximising model was calibrated for a given suboptimal situation, then they used the linear programming model {(64), (65), (66), (67) and (68)} to find the optimal solution. They also checked that the linear programming estimates of T_{ij}^{kw} are the limits of the entropy maximising model estimates as μ is increased [12].

Inter-regional Commodity Flows and Energy Flows in an Urban Space Economy

Entropy maximising techniques can be used in urban or regional economic forecasting. Models can be built to accommodate price increases and energy shortages to meet the prevailing energy constraints.

Denote a set of regions by $i, j, k, \ldots$ and the economic goods classified into commodity goods by $m, n, p, \ldots$. Also define

x_{ij}^m = the total flow of commodity m from region i to region j ,

c_{ij}^m = the mean cost of transporting a unit of commodity m from i to j,

X_i^m = the total production of commodity m in region i ,

Y_i^m = the total consumption of commodity m in region i .

In exactly analogous way to the trip-distribution case, a production-attraction constrained model can be derived for the inter-regional flows of commodities. The constraints are:

$$\sum_j x_{ij}^m = X_i^m \quad , \tag{69}$$

and

$$\sum_i x_{ij}^m = X_j^m \quad . \tag{70}$$

The third constraint takes the form

$$\sum_i \sum_j x_{ij}^m c_{ij}^m = C^m \quad , \tag{71}$$

where C^m is the total expenditure in the system on transporting commodity m . Defining a Lagrangian and proceeding in the usual way, the solution is

$$x_{ij}^m = e^{-(\lambda_i^{(1)m} + \lambda_j^{(2)m} + \mu^m c_{ij})} \quad , \tag{72}$$

where $\lambda_i^{(1)m}$, $\lambda_j^{(2)m}$ and μ^m are the Lagrange multipliers associated with the constraints (69), (70) and (71). The parameter μ^m has to be estimated (for each commodity group m) by calibration.

If we let

$$e^{-\lambda_i^{(1)m}} = A_i^m X_i^m , \tag{73}$$

and

$$e^{-\lambda_j^{(2)m}} = B_j^m Y_j^m , \tag{74}$$

and substitute in (69) and (70), equation (72) becomes

$$x_{ij}^m = A_i^m B_j^m X_i^m Y_j^m e^{-\mu^m c_{ij}^m} , \tag{75}$$

where

$$A_i^m = \frac{1}{\sum_j B_j^m Y_j^m e^{-\mu^m c_{ij}^m}} , \tag{76}$$

and

$$B_j^m = \frac{1}{\sum_i A_i^m X_i^m e^{-\mu^m c_{ij}^m}} . \tag{77}$$

Equations (75), (76) and (77) are the production-attraction constrained model for inter-regional flows of commodities. Production-constrained and attraction constrained sub-models of this model have been derived by Wilson [11], and interpretations similar to those in the trip-distribution case adopted.

Models of energy and materials flows in an urban space economy would be of importance in making regional economic forecasts. They could also be a contribution to the study of changing energy situations.

Wilson et al [1] built models for urban energy sectors by constraining the flows and then using the entropy-maximising technique to obtain most probable internally consistent estimates of the variables. For example, for energy sector changes that affect transport patterns, energy usage and its cost for various pricing policies (c_{ij}'s) and distributions of land use (O_i's and D_j's) could be explored. This is usually achieved by the introduction of additional constraint equations.

6. CONCLUDING REMARKS

From what has been outlined in the above sections, it can be seen that entropy plays an important role in the formulation of mathematical models for environmental studies. One would, however, make some observations: these will be related with dynamics in the modelling field, model-consistency, and application to towns like Lusaka.

The models that have been discussed have been taken to be describing an equilibrium situation, that is, parameters have been assumed to be constant with time. This is never the case since urban patterns change with time. Then, perhaps, an important research task would be parameter forecasting.

It is important that interaction models should be internally consistent. In entropy maximising techniques, this is achieved by specifying the main interaction variables. Wilson [9] attained high levels of resolution while still being consistent. If, for instance, the categorisation by person type was not made, one would be faced with a situation where non-car owners are allocated car trips. While entropy-maximising models are internally consistent, the number of parameters to be estimated also increases to uncomfortable proportions with greater disaggregation.

Although these mathematical models have been applied to large cities like Leeds and London, they can be applied to towns like Lusaka since there is also need to plan for the future. Of course fewer people, smaller areas and smaller income groups would be involved, thereby allowing for high levels of resolution to be attained.

A traffic density count was carried out in Lusaka by Emenalo [2]. Emenalo states in the report that "A number of road junctions in Lusaka are unnecessarily staggered. Such junctions have been found to be accident black spots...", and that "Filter lanes are almost non-existent in Lusaka, particularly at junctions where their presence would undoubtedly reduce traffic congestion..." . This shows lack of good planning, and the models discussed in this dissertation would be of relevance.

REFERENCES

[1] CRIPPS E.L., MAcGILL S.M., AND WILSON A.G. (1974): Energy and Materials Flows in the Urban Space Economy. Transportation Research 8, pp 293-305.

[2] EMENALO, S.I. (1970): Road Traffic Prob lems in Zambia: Interim Report on Density Count. University of Zambia Publication.

[3] EVANS, E.L. (1970): Some Properties of Trip Distribution Methods. Transportation Research, 4, pp 19-36.

[4] HOBSON, A. : Concepts in Statistical Mechanics. Gordon and Breach Science Publishers, 1971.

[5] SENIOR, M.L. AND WILSON, A.G. (1974): Explorations and Synthesis of Linear Programming and Spatial Interaction Models of Residential Location. Geographical Analysis, 6, pp 209-238.

[6] SHANNON, C.E. AND WEAVER, W.: The Mathematical Theory of Communications. University of Illinois Press 1964.

[7] TOLMAN, R.C.: Principles of Statistical Mechanics. Clarendon Press , Oxford, 1938.

[8] WILSON, A.G. (1968): Notes on some Concepts in Social Physics, Regional Science Association Papers XXII.

9] WILSON, A.G. (1969): The Use of Entropy Maximising Models in the Theory of Trip Distribution, Mode Split and Route Split. Journal of Transprt Economics and Policy, 3, pp 79-85.

0] WILSON, A.G. (1970): Advances and Problems in Distribution Modelling.
Transportation Research, 4, pp 1-18.

1] WILSON, A.G. (1970): Inter-regional Commodity Flows: Entropy Maximising Approaches.
Geographical Analysis, 2, pp 255- 282.

2] WILSON A.G. AND SENIOR, M.L. (1974): Some Relationships Between Entropy Maximising Models, Mathematical Programming Models, and their Duals.
Journal of Regional Science, 4, No. 2.