SUBJECTIVE SAMPLING APPROACHES TO RESOURCE ESTIMATION

Gregory B. Baecher

June 1975

Subjective Sampling Approaches to Resource Estimation

Gregory B. Baecher

Abstract

This paper suggests deficiencies in present sampling approaches to regional resource estimation, and ways in which these deficiencies might be remedied. General approaches to resource estimation are discussed, as are requirements which well conceived approaches should satisfy. Using presently available theory, a comprehensive sampling approach to estimation can be formulated. The results of such an analysis are directly incorporable into decisions concerning exploration strategy optimization. However, further computational and experimental work are required before this approach is operational.

I. Introduction

Resource estimation techniques can be broadly grouped into two classes: macroanalytic approaches which model empirical relationships in aggregated discovery or production data, and microanalytic approaches which model structural relationships in the exploration process. Perhaps the most well known examples of each of these are Hubbert [14] and Allais [1].

This paper addresses microanalytic approaches. In particular, it sets about broadening present sampling theory techniques to encompass more of what we know about

the exploration process, and to include prior geological
opinion. This broadening is seen as necessary if estimates
based on microanalytic methods are to be comprehensive and
valid.

The main purpose of the paper is not to mathematically
solve formulae associated with the broadening, but to
indicate directions toward which continuing work should be
moving.

II. Macroanalytic versus Microanalytic Approaches

Macroanalytic models, which in essence are trend
extrapolation procedures, assume an unspecified "uniformity-
of-nature." They assume that exploration and production
operate within a fixed (or at most, gradually changing)
environment which leads to aggregate behavior according to
simple relationships among important variables. Taking
this to be true, empirically fitted relationships may be
extrapolated into the future, either in time or along
some other dimension (e.g., cumulative drilling length).
Macroanalytic approaches do not use structural relation-
ships among facets of exploration and production, and
lump together economic, geological, and exploratory
variables.

Arguments for and against macroanalytic approaches
appear in the geologic literature (Ryan [26], Hubbert [14],
Moore [18]) as well as in the literature of other disciplines
where similar tools are used for estimation or forecasting
(e.g., economics). Specifically, two properties limit their

usefulness for resource estimation. First, they lead to deterministic predictions, the uncertainty of which is difficult to judge (e.g., changing from one family of curves to another, or from one method of fitting to another, drastically changes estimates—Cf., Hubbert [14], Moore [18]). Second, they depend on a substantial history of discovery and production. While this history exists for areas like the United States, for sparsely explored areas the analysis often begins by predicting total resources some other way, and then calculates time streams of production (Hubbert [14]).

Microanalytic approaches also suffer drawbacks, which again are generic to the approach and not limited to resource estimation. First, they require detailed data on a region by region basis of geologic and geometric properties, numbers and sizes of discoveries, and amounts and patterns of exploration allocations. Second, they require orders of magnitude more computation effort than macroanalytic approaches, as gross estimates are formed by first making regional estimates and then aggregating. These requirements make microanalytic approaches difficult and laborious to apply. Third, although not necessarily a shortcoming, microanalytic approaches do not account for economic or production factors. They deal purely with geological and statistical variables. Economic variables must be considered separately using the geological/statistical analysis as input (e.g., MacAvoy and Pindyck [17]).

On the other hand, microanalytic approaches have four
very favorable properties which recommend them from the
present perspective:

1) they allow inclusion of geological input on a
   regional basis;

2) they may be applied to regions which have been
   only sparsely explored;

3) they often allow quantification of uncertainty; and

4) their output can be readily incorporated in strategy
   optimization for local or regional exploration.

III. Microanalytic Models

Microanalytic approaches proceed by making resource
estimates for small regions which are assumed geologically
homogeneous, then aggregating over all regions.
As the aggregation is straightforward, attention is drawn
to making estimates for each region. In a traditional,
judgmental way this has always been done by exploration
geologists. Based on experience, geologists subjectively
judge the similarity of the region to better known regions,
and in combination with geological theory make predictions
of resources (e.g., upper and lower bounds). This is a
very basic microanalytic approach, and is the approach
which Harris [11] attempts to quantify.

A second approach is to correlate geological variables
with resources either by regression or factor analysis
(Harris [12], DeGeoffroy and Wingall [4], DeGeoffroy and
Wu [5]). This is a straightforward approach with which

there is experience in many applications. However, it suffers well-known limitations in that it is a correlation and not a causal model. Factors which are highly correlated with mineralization or deposition in one context are not necessarily those which would be correlated with it in others. As these methods are normally applied to known deposits rather than resources, regression and factor analysis confound geological and non-geological variables. This leads to the not too surprising result of Griffiths and Singer [10] that "mineral potential" is most highly correlated with degree of development.

A third approach treats estimation as a problem of inference from sampling. The size distribution and spatial dispersion of deposits are modeled by families of probability functions, and parameters for these distributions are estimated assuming known deposits to be a probability sample of the total in situ population (Allais [1], Uhler and Bradley [32], Kaufman [15], Slichter [30], Griffiths [9]). Total resource estimates are made by evaluating the random sum

$$R = a_1 + a_2 + a_3 + \cdots + a_N \tag{1}$$

in which a_i is a random variable drawn from the distribution of deposit sizes, and N is a r.v. representing the number of deposits within the region (Uhler and Bradley [32]).

Criticism of the sampling approach has primarily been based on the observation that known deposits are not a simple random sample of in situ deposits, the necessity of choosing families of distributions to model the size distribution and spatial dispersion of deposits, and lack of geological input in the model.

The observation that known deposits are not a simple random sample of in situ deposits is not so much an argument against a sampling approach as an argument against uncritical application of that approach. For example, Kaufman [15] has presented a more rigorous analysis of the sampling approach in which in situ deposits are treated as a finite random sample from some "super-population" (see also Ericson [7]). Then, the parameters of that super-population are estimated by assuming known deposits to be a sample of the in situ population selected with probability proportional to size and without replacement (Figure 1). Very different estimates of super-population parameters are obtained using this assumption from those using the simple random assumption. A point we will return to in Section 3 is that similar special considerations must be made in estimating parameters of the spatial dispersion function, in particular, that the probability of finding n deposits within a subregion is nonlinearly related to the amount of exploration effort allocated to that subregion and to the distribution of deposit sizes.

The problem of selecting a family of distributions to model the super-population and spatial dispersion of deposits is common to all analyses (in that mathematically simple functions must always be chosen somehow). In the present context, however, there is considerable empirical evidence to suggest that lognormal super-populations do accurately model geometric properties of many geological populations (Slichter [30]), and that the negative binomial distribution may adequately model spatial dispersion, though the latter point is not as clear as the former. We will return to this spatial model in Section 3. Furthermore, one suspects (see, e.g., Uhler and Bradley [32]) that the total resource estimates are fairly robust to changes in the form of the super-population, and even more so to the form of spatial dispersion.

An interesting direction of future work would be to quantitatively evaluate the sensitivity of resource estimates to the form of these distributions. A refinement following such analysis may be to form so-called composite Bayesian distributions as suggested by Wood [34] and Box and Tiao [3], in which distribution models are themselves random variables.

The most important criticism of sampling approaches is that they generally neglect prior geological information. This is certainly true of the "frequentist" approaches to inference, and even the Bayesian analyses have remained tied to "unintormed priors." In making estimates of natural resources we have considerably more information available than merely

the number and sizes of already discovered deposits. This prior information comes from regional geology, experience in similar regions, and basic concepts of geological processes. Comprehensive estimates must account for this information. It is only when the available data set is so large that inferences become insensitive to prior information that the latter can be neglected. Given the small amount of information which comes from finding or not finding deposits (relative to the inferences about regional geology and structure which are made), this is seldom the case. Given human biases toward neglecting prior information in the face of new, "hard" data (Tversky [31]) inclusion of geological information must be explicit.

IV. Requirements for a Comprehensive Sampling Approach

To this point we have discussed macro- and microanalytic approaches to resource estimates, and indicated advantages and disadvantages of each. From this discussion it seems apparent that sampling approaches offer a methodological framework within which a comprehensive and realistic model of exploration and estimation might be developed. We now turn toward necessary modifications of the sampling approach.

Two requirements which present sampling approaches do not entirely satisfy, but which they must to be comprehensive and realistic are that:

a) prior geological information and opinion be accounted for;

b)   the real likelihood of deposits being discovered

      be reflected.

Logically, these facets of inference are separable and may be combined by Bayes' Theorem,

$$f'(\underline{\theta},\Omega|data) \propto f^{o}(\underline{\theta},\underline{\Omega}) \ L(data|\underline{\theta},\Omega) \quad . \tag{2}$$

Here, $\underline{\theta}$ and Ω are taken to be geological parameters descri-bing the size and spatial distribution of deposits; $f^{o}(\underline{\theta},\underline{\Omega})$ and $f'(\underline{\theta},\Omega/data)$ are the prior and posterior prob-ability distribution of the parameters, respectively; and $L(data/\underline{\theta},\underline{\Omega})$ is the likelihood of <u>observing the data</u> were $\underline{\theta}$ and Ω the true parametric value. Prior geological infor-mation is contained in $f^{o}(\underline{\theta},\underline{\Omega})$; characteristics of the exploration process are contained in $L(data/\underline{\theta},\underline{\Omega})$.

4.1 <u>Prior Information and Subjectivity</u>

One enters nearly all inferential situations with some prior information or suspicions. A region seems favorable for exploration because it is similar to known areas of deposition or because it has geological properties associ-ated with deposition. However, each individual has dif-ferent experiences and concepts of geology and thus assesses favorability differently. This is the traditional role of the exploration geologist. Geological structures are highly complex, and <u>comparatively</u> few observations are made in exploration. Therefore, experience and judgment are im-portant. This is why geologists are called upon to make

resource estimates rather than other people (see Robinson
[24], for an illustration of the importance of subjective
concepts in interpreting exploration data).

A geologist considers the results of exploration in
the context of his prior feelings. To the extent the two
are consistent he gives more or less credibility to his
feelings. However, this inferential process, and thus
exploration as a whole, is purely subjective. Thus
exploration cannot be adequately modeled without intro-
ducing the concept of subjective probability (Baecher [2]).
The useable results of exploration are hypotheses.
These hypotheses arise subjectively and are given credence
subjectively; "hard" data only enter in modifying the
degree-of-credibility given to hypotheses. Uncertainties
associated with exploration are those associated with
hypotheses, so uncertainties, too, are necessarily sub-
jective. Only when the amount of data becomes so large
that inferences cease to be affected by prior feelings can
exploration be spoken of as "objective." This occurs for
only the most intensively explored regions.

So, the sampling approach we would like to develop
must be based on subjective probability. This is not unique
to the exploration literature, although a thorough attempt
at a rigorous fundamentally subjective approach may be.[1]

---

[1]Grayson's [8] well-known application of statistical
decision theory to oil and gas drilling decisions is, of
course, a rigorous and early application of subjective
probability to geological exploration. However, for what-
ever reasons, subjectivism has never been adopted by
"geostatisticians" and thus the resource estimation liter-
ature remains non-subjectivst and (with the exception of
Kaufman) non-Bayesian.

Kaufman [15] bases his analysis on a Bayesian approach,
but does not use geological information to assess
priors (adopting "uninformed" priors instead). Harris [11]
and Harris, et al. [13] use subjective probability in a
one-step procedure for making resource estimates without
exploration data (i.e., using only geological maps).
However, this is a degenerate case of resource estimation,
and they seem to use subjective probability merely as a
pragmatic tool when other data are not available.

Assessing Subjective Probabilities

Applicability of subjectivist theory rests ultimately
on our ability to adequately assess probability distri-
butions. Adequacy here means the ability to quantify an
individual's true personal feelings in a probability measure.
There is not room here to review the literature on behavioral
decision theory and quantification of subjective probabili-
ties. However, this work is extensive and rather consistent.
Feelings can be reliably quantified if a careful, rigorously
based technique is employed. People do exhibit bias in
quantifying their feelings, (Tversky [31]) but these biases
may not be great. Individuals may exhibit consistent con-
servative biases in updating their prior feelings by sample
data (Edwards [6]), but in highly complicated, real
problems this conservatism seems to diminish or even dis-
appear (Winkler and Murphy [33]). In some meteorological
experiments, experts' measured, subjective probabilities
have been shown to be better forecasters of natural

occurrences than structural models (Murphy and Winkler [22]). In short, we can adequately assess subjective probabilities, but these assessments should be carefully made in the context of past research. As with any technology, haphazard application leads to unreliable results.

Coalescing Geological Opinion

Adopting a subjectivist philosophy of course leads to the problem of differing expert opinion, and speaking of "good" and "bad" assessors ceases to make sense. Probabilities reflect only individual feelings, which in turn may not reflect reality. These differences are no surprise, however, as the literature contains wildly differing resource estimates already, and policy makers have always had to deal with differing expert opinions.

The traditional way to coalesce differing subjective probabilities has been the Delphi method, which is a discussion and averaging process. This procedure has received considerable criticism, but is widely used (Pill [23]). Actually, it is more consensus seeking than a true coalescence. Harris [11] uses this approach in his mineral potential estimates of Sonora.

A more rigorous method based entirely on Bayesian philosophy has been recently proposed by Morris [19, 20], and this approach could be adopted for coalescing geological opinion. Assume that the ultimate policy analyst, can himself assign some prior subjective distribution to the extent of deposition or mineralization. Let these estimates be

expressed in terms of two sets of parameters which correspond respectively to the size distribution of deposits $(\underline{\Theta})$, and to spatial dispersion $(\underline{\Omega})$. These prior probabilities could be taken as uniform. Opinion is taken individually from several geologists in the same terms, that is, in the form of probability distributions on the parameters $\underline{\Theta}$ and $\underline{\Omega}$. To the analyst or policy maker these probability distributions (representing expert opinion) are <u>information</u> and he may coalesce them by the normal Bayesian argument, using his own feelings as a prior,

$$f'(\underline{\Theta},\underline{\Omega}|\text{experts' opinion})$$

$$\propto f^{o}(\underline{\Theta},\underline{\Omega}) \ L(\text{experts' opinion}|\underline{\Theta},\underline{\Omega}) \ . \quad (3)$$

This formulation offers a rigorous relationship for coalescing expert opinion. The difficulties of evaluating "credibility" of experts are concentrated (some might say transferred) to developing a likelihood function for their opinion conditioned on what the actual parametric values $\underline{\Theta}$ and $\underline{\Omega}$ might be. While this is straightforward, it becomes untidy when the likelihoods of individual experts' opinions are not independent.

But how can the likelihood function be estimated? As Morris argues, no matter how one proceeds with a statistical analysis, likelihood functions are always established subjectively. For convenience, we may assign families of distributions to those as we do to other things (e.g., a normal

likelihood) but always this is judgmentally done. Just as
we assess subjective probability, so also we may assess
likelihood functions based on the policy-analysts' or de-
cision-maker's feelings relative to his experts' credibility.
This reflects the central argument in favor of all
quantitative decision analysis: quantitative analysis does
not make decisions for the decision maker, rather it allows
him to decompose a decision (or estimation), treat each
part in isolation, then reaggregate in a logically consis-
tent manner to draw deductive conclusions. Always, the
conclusion drawn rests on the judgment of the person who
draws it. To deny this is misleading.

A strength of this approach is that it allows the
analyst also to establish the expected value of expert
opinion (or the marginal expected value of an additional
opinion). This process is established exactly as the
"expected value of sample information" is evaluated in
any Bayesian Decision Theoretic application.

A Proposal for Including Geological Opinion in Resource Estimates

Entering a new estimation task there are four types of
prior information to be included: individual experience,
documented experience, geological theory, and local
characteristics. Were there only documented experience
and local conditions, priors could be generated by regression
or related techniques. However, individual experience and
theory serve to modify direct correlations with the "hard"

data of previously explored areas by degrees to which the region under consideration is or is not similar to previous areas, and the ways in which it seems anomalous in terms of basic geological processes. In combining these sources of information the geologist functions somewhat as a subjective information processor (Figure 2).

The approach proposed is that each geologist be <u>given</u> information in the form of geological properties and estimates of $\underline{\theta}$ and $\underline{\Omega}$ for grossly similar regions in which more extensive exploration has been conducted, and local characteristics of the region in question. Then through a process of careful questioning and gaming directly assess his feelings about possible values of $\underline{\theta}_0$ and $\underline{\Omega}_0$, the local parameters, in terms of probability measures. This process might be extended by preconditioning data from other regions in terms of local characteristics (i.e., regression or factor analyses applied to the new region). In this way, each expert bases his judgment primarily upon the same hard data set, and incorporates his past individual experience and concepts of geological processes purely subjectively.

As Morris points out, it is not a simple task to ascertain the independence of expert opinion. If the opinion is independent, the likelihood function of equation (3) reduces to the simple multiplicative form of the marginal likelihoods; but if it does not, interdependencies must be modelled, and these may have complex and non-obvious forms.

In particular, if experts base their judgments partially upon the same data, then their opinions are not independent. Proceeding as outlined above, however, mitigates this dependency by forming opinions which are conditioned on the data set, and thus may be conditionally independent which would allow a simple multiplicative form.

Mitigating the problem of dependence caused by similar geological theory is not so easily achieved, and indeed will require further attention to the design of assessment schemes.

The second step of the process is coalescing opinion. How can likelihoods of geologists' opinion be generated? Currently this problem is difficult to treat except in simplistic ways, but the theoretical base of this approach is only now expanding (e.g., Morris [20]). As a first approximation one can assume that the likelihood of a prediction is related only to the absolute value of the discrepancy from the true parametric value. That is, that experts' opinions are unbiased and that error is symmetric about true values. If one assumes a simple analytical function, e.g., a normal distribution, to represent this error, then the variance of that distribution is a sufficient description of expert credibility. It would fall to the analyst or policy maker to subjectively decide upon values of this variance (i.e., "credibility") for each expert he consults--but this is always the task of the analyst whether he achieves it quantitatively or qualitatively.

Symbolically, this analysis is of the form

$$f^o[\underline{\theta}_o, \underline{\Omega}_o | \{f_i(\underline{\theta}, \underline{\Omega})\}] \propto f^a[\underline{\theta}_o, \underline{\Omega}_o]$$

$$\times \prod_{i=1}^{z} \int_{\underline{\theta}} \int_{\underline{\Omega}} f_i(\underline{\theta}, \underline{\Omega}) \; f_n(\underline{\theta}, \underline{\Omega} | \underline{\theta}_o, \underline{\Omega}_o, \underline{\underline{\Sigma}}) \; d\underline{\theta}_o d\underline{\Omega}_o \quad . \tag{4}$$

Where $f^o(\cdot)$ is the probability distribution used as a prior in subsequent resource estimates, $f^a(\cdot)$ is the analysts prediction of the parameters (which might be uniform), $f_i(\cdot)$ is the ith geologist's prediction, and $f_n(\cdot | \underline{\theta}_o, \underline{\Omega}_o, \underline{\underline{\Sigma}})$ is the normal distribution (in this case the likelihood) with mean $\underline{\theta}_o, \underline{\Omega}_o$ (i.e., the assumed true values) and variance matrix $\underline{\underline{\Sigma}}$. As a first approximation, it seems reasonable to assume that errors in the estimate of $\underline{\theta}$, the parameters of the size distribution, and $\underline{\Omega}$, the parameters of spatial dispersion are independent. So,

$$\underline{\underline{\Sigma}}_i = \begin{bmatrix} \sigma^2_{i\theta} & 0 \\ 0 & \sigma^2_{i\Omega} \end{bmatrix} , \tag{5}$$

in which $\sigma_{i\theta}$ is the credibility assigned to geologist i's estimate of $\underline{\theta}$, and $\sigma_{i\Omega}$ is the credibility assigned to his estimate of $\underline{\Omega}$.

The approach just described, clearly, is very rough. Considerable effort, and in particular attempts to apply such methodologies, would need to be invested before a workable and practical procedure could be developed. Nevertheless, an approach somewhat of the type outlined is needed to analytically include geological opinion within the context

of regional resource estimation. Ignoring this prior
information leads to estimates which are not comprehensive,
overly diffuse, and possible erroneous.

4.2 Likelihood Function

In the Bayesian scheme, equation (2), characteristics of the
sampling plan are entirely contained within the likelihood
function. This is the probability of obtaining the sample
actually observed--that is, the deposits actually discovered
--conditioned on values of the parameters $\underline{\theta}$ and $\underline{\Omega}$. This
probability may or may not depend on the order of discovery.

For simple random sampling each observation is assumed
independent, and their ordering unimportant. If deposits
of size x_i are discovered in this way, their likelihood is

$$L(\underline{a}|\underline{\theta}) = p(a_n|a_1, \ldots, a_{n-1}, \underline{\theta}) \; p(a_{n-1}|a_1, \ldots, a_{n-2}, \underline{\theta})$$

$$\ldots p(a_2|a_1, \underline{\theta}) \; p(a_1|\underline{\theta}) \quad , \tag{6}$$

$$= \prod_{i=1}^{n} t(a|\underline{\theta}) \quad ,$$

where $f(a|\underline{\theta})$ is the distribution of deposit sizes from which
discoveries are made.

As Kaufman [15] points out, however, discoveries of
mineral deposits do not follow a simple-random process.
First, the total population of in situ deposits is finite;
and second, larger deposits have a greater probability of
being found than smaller ones. Once a deposit is found it
is "removed" from those which might still be found, and

thus sampling is "without replacement." This means that the order of discovery is important.

Kaufman assumes that deposits appear in the sample with probability proportional to the ratio of their size to the cumulative size of still undiscovered deposits. This is the probability relative to other deposits appearing, or the probability conditioned on a discovery. He also postulates that in situ deposits be considered a simple-random sample from some infinite super-population, $f_s(x|\underline{\theta})$, then inters values of the parameters of that distribution. Considering equation (1) once again, this approach allows inferences on the distribution of the random variables x_i in the resource estimate, and also inferences about the sum of undiscovered sizes. It does not allow direct inferences of the in situ number, N.

While this procedure offers an approach to estimating total resources, it does not make use of all available information, and does not yield spatial characteristics which might be used in optimizing future exploration strategies. However, it may be expanded to include the likelihood of numbers of deposits being discovered and the non-uniform geographic distribution of exploration, and thus to overcome these objections.

Spatial Dispersion Function

The spatial dispersion of mineral deposits is most often treated as a point process in two dimensions.[2] Parameters of the spatial dispersion model are then estimated by dividing the geographic region into quadrats and fitting curves to the distribution of numbers of deposits per quadrat.

Empirical data displays more clustering than the Poisson model would predict, thus other models have been considered and at present there seems widespread satisfaction with the negative binomial model (DeGeoffroy and Wu [5], Griffiths [9], Uhler and Bradley [32]--other models are discussed in Rogers [25]). Among the few criticisms of the negative binomial is that it tends to underestimate the frequency of quadrats with high numbers of deposits (Kaufman and Bradley [16]) (Figure 3).

Previous work typically assumes the number of known deposits per quadrat to be mutually independent samples from the spatial dispersion process; thus,

$$L(n_1,\ldots,n_q \mid \underline{\Omega}_o) = \prod_{i=1}^{q} P_{nb}(n_i \mid \underline{\Omega}_o) \quad , \tag{7}$$

where q is the number of quadrats, n_i the number of known deposits in quadrat i, and $\underline{\Omega}$ the parameters of the spatial

---

[2]Kaufman and Bradley's [16] random-walk simulation is one of the few exceptions.

dispersion process (whose values are to be inferred). This procedure leads to results which are difficult to interpret for the following reasons:

1) Known numbers of deposits are not samples from the spatial process $p(N|\underline{\Omega})$ but are lower bounds on the actual number in a quadrat.

2) If the analysis is restricted to intensively explored quadrats, which would yield truer samples of $p(N|\underline{\Omega})$, the sample of quadrats is biased toward greater density (i.e., the most intensively explored quadrats are also the ones with the most extensive mineralization or deposition).

3) If very sparsely explored quadrats are included, the sample is biased toward low numbers per quadrat; the probability of discovering in situ deposits in these quadrats is small.

This approach clearly leads to incorrect estimates.

Search Effort

The number of deposits found in exploration obviously depends on the amount and spatial allocation of search effort. If this effort is non-uniformly distributed geographically, then the probability of discovery is non-uniform also. Although this principle is intuitively clear, it may explain certain anomalies in resource modelling, and may lead to mitigation of the three objections just mentioned.

Assume temporarily that deposits were actually dispersed according to a negative binomial process. Then let one deposit be found in some quadrat, c, as shown in Figure 4. As deposit locations are positively correlated, this increases the favorability of quadrat c for containing additional deposits. That is, the probability of c containing at least one more deposit is increased from 0.19 to 0.52 (using DeGeoffroy and Wu's parameters). Therefore, an optimal exploration strategy would be to allocate more effort to exploring quadrat c than other quadrats. Since this process feeds back upon itself as more discoveries are made, high n quadrats appear in observations with probability disproportionately higher than their frequency in situ. Thus the objection of Kaufman and Bradley may only reflect non-uniform exploration.

Returning to equation (7), one sees that the likelihood is not merely the spatial dispersion model, but must be modified by the probability of finding in situ deposits. We will call this latter relation the detection function. The detection function has the property that when there is no exploration effort ($\psi = 0$) the probability of a discovery is zero ($p(n = 0) = 1.0$), and as $\psi \rightarrow \infty$, $p(n = N) \rightarrow 1.0$. Here n and N are the number of discovered deposits and the total number of in situ deposits, respectively. So, as is intuitively clear, the probability of discovering deposits within a quadrat depends on the number of deposits present and the effort exerted to find them.

Form of the Detection Function

While the detection function begins at zero and reaches an asymptote of 1.0, its exact form depends on the strategy of allocating search effort and the distribution of deposit sizes.

Consider a quadrat of area A which contains a single deposit of area a. If ψ units of search effort are randomly allocated to points within the quadrat, the probability of finding the deposit is (Figure 5),

$$\Pr(\text{find}|\psi) = 1 - (1 - a/A)^{\psi}$$

$$\cong 1 - e^{-\psi(a/A)} \qquad a/A < 0.1 \quad . \qquad (8)$$

If a systematic allocation is used (i.e., a grid), then p(find) depends both on the target and grid geometries, as illustrated in Figures 6a to 6d. Similar curves can be generated for other systematic allocations (e.g., geophysical methods) or for "optimal search" when prior locations probabilities can be specified (Morse [21]).

Without detailed information on the way exploration has been conducted, there is no way to precisely reconstruct the detection function. Therefore, in making resource estimates we must make assumptions on its shape. On one hand, exploration may be viewed as the uncoordinated effort of many separate decision makers. If this is so, then a random model seems appropriate. On the other hand,

exploration may be carried out by one decision maker as in the case of a government ministry or large corporation. Were this the case, then a purely systematic model might be appropriate. Both are crude approximations, but perhaps satisfactory first attempts.

Both random and grid search can be approximated by an exponential detection function of the form[3]

$$p(\text{find}|\psi, a) = 1 - e^{-k\psi a/A} , \tag{9}$$

in which k is a constant. This can be modified for uncertainty in deposit size in the normal way,

$$p(\text{find}|\psi, \underline{\theta}) = \int_a p(\text{find}|\psi, a) \, f(a|\underline{\theta}) \, da . \tag{10}$$

---

[3]Ryan's [26, 27] deterministic model of discovery within a play is of this form, though he does not directly treat it as a detection function. In his model cumulative-newfield-wildcats is used as a measure of ψ, and he introduces a constant multiplied by ψ to account for "geological knowledge." To find the regional resource he equates rate of new discoveries to the product of resource and detection function.

$$R = U_\infty \left[1 - e^{-\beta k w} \right] ,$$

where

R = rate of discovery,

U_∞ = total resource,

β, k = constants,

w = cumulative new wildcats.

This deterministic model closely fits empirical rates of discovery within individual plays, and thus adds credibility to the random exploration model. However, his equation has three adjustable parameters and thus is flexible.

To form the likelihood function for inferring values
of the spatial parameters, $\underline{\Omega}$, the number of discoveries
must be related to the number of in situ deposits by an
equation of the form

$$p(n|\psi,\Omega) = \sum_N p(n|N,\psi)\, p(N|\psi) \quad . \tag{11}$$

Here, $p(n|N,\psi)$ is a modification of the detection function
to account for multiple deposits, and $p(N|\underline{\Omega})$ is the spatial
dispersion process. Unfortunately, $p(n|N,\psi)$ is not a simple
relationship.

Let n deposits be found in a particular quadrat
in the order

$$a_1,\ a_2,\ \cdots,\ a_n \quad ,$$

with

$$\psi_1,\ \psi_2,\ \cdots,\ \psi_n$$

increments of exploration effort, respectively. Given that
the first $j-1$ of these have been found, the probability of
finding the jth with one additional quantum of effort is

$$p(\text{find}|a_1,\ldots,a_{j-1},\psi=1) = p_j = \frac{\displaystyle\sum_{i=j}^{n} a_i - \sum_{i=n}^{N} a_i}{A - \displaystyle\sum_{i=1}^{j-1} a_i} \quad ; \tag{12}$$

and the probability of having discovered the jth deposit
with the increment of effort, ψ_j, is

$$p(a_j|a_1,\ldots,a_{j-1},\underline{\Theta},N) = \text{Pr(find on } j^{th} \text{ quantum})$$

$$\times \text{ Pr}(a_j \text{ find}) \times f(a_j|\underline{\Theta}) \text{ ,}$$

$$= \exp\{-k\psi_j p_j\} p_j \frac{a_j}{\sum\limits_{i=j}^{n} a_i + \sum\limits_{i=n}^{N} a_i} f(a_j|\Theta) \quad P_n^N \text{ .} \tag{13}$$

As N is a random variable with parameters $\underline{\Omega}$, this becomes

$$p(a_j|a_1,\ldots,a_{j-1},\underline{\Theta},\underline{\Omega})$$

$$= \sum_N \exp\{-k\psi_j p_j\} p_j \frac{a_j}{\sum\limits_{i=j}^{n} a_i + \sum\limits_{i=n}^{N} a_i} f(a_j|\Theta) \; p(N|\underline{\Omega}) \; P_n^N \text{ ,} \tag{14}$$

in which the term $S = \sum\limits_{i=n}^{N} a_i$, the sum of undiscovered deposits,

is an uncertain quantity depending both on $\underline{\Theta}$ and $\underline{\Omega}$. The

likelihood of discoveries then is

$$L(a_1,\ldots,a_n|\underline{\Theta},\underline{\Omega})$$

$$= \sum_{j=1}^{n} \int_S \sum_N \exp\{-k\psi_j p_j\} p_j \frac{a_j}{\sum\limits_{i=j}^{N} a_i + S} f(a_j|\underline{\Theta})$$

$$\times p(N|\underline{\Omega}) \; f(S|\underline{\Theta},\underline{\Omega}) \; P_n^N \; dS \text{ .} \tag{15}$$

Clearly, this equation is difficult to deal with, although as Kaufman has done, this might be approached by Monte Carlo simulation. It does account for exploration effort, however, and conceptually at least allows inferences to be drawn about the spatial dispersion of deposits.

The point of this short discussion is that inferences about the number of deposits in a quadrat or region (and thus about their spatial dispersion and the total amount of resources) must account for how and how hard they were looked for. Further, inferences about spatial dispersion are not independent of inferences about size distribution; the simple-random sampling model is not satisfactory for this purpose.

V. Conclusions

This paper has discussed the place of sampling approaches to resource estimation in a broad context, and it has attempted to indicate that sampling approaches could lead to a more comprehensive analysis than is currently employed. Specifically, discussion has concentrated on three points about exploration and inferences drawn from it:

a) Geological exploration is fundamentally and necessarily a subjective undertaking; prior judgment of geologists based on findings in other regions and on concepts of geological processes must be included.

b) The analytical methods for including geological opinion from multiple experts must be theoretically rigorous and reflect current knowledge of probability assessment, judgmental biases, and subjective information processing.

c) The procedure used for modifying prior opinion by the local results of exploration should include consideration of exploration effort and its allocation through some detection function.
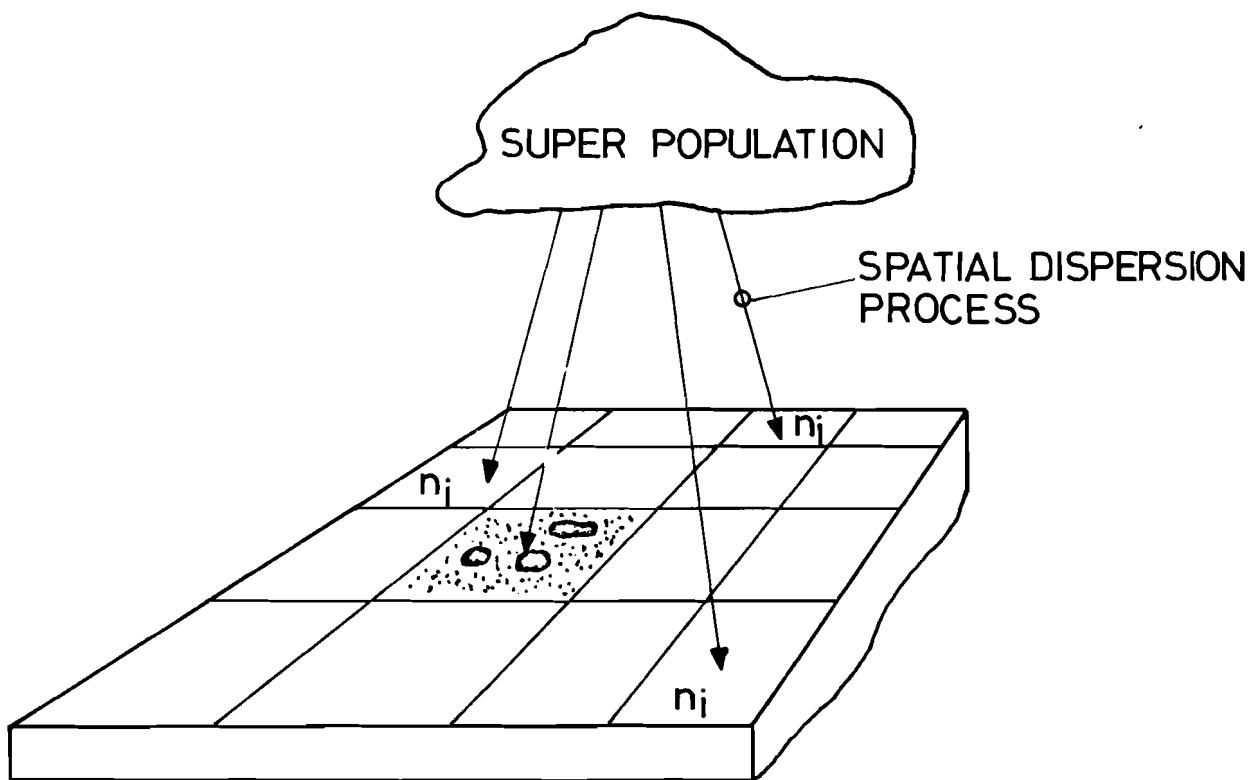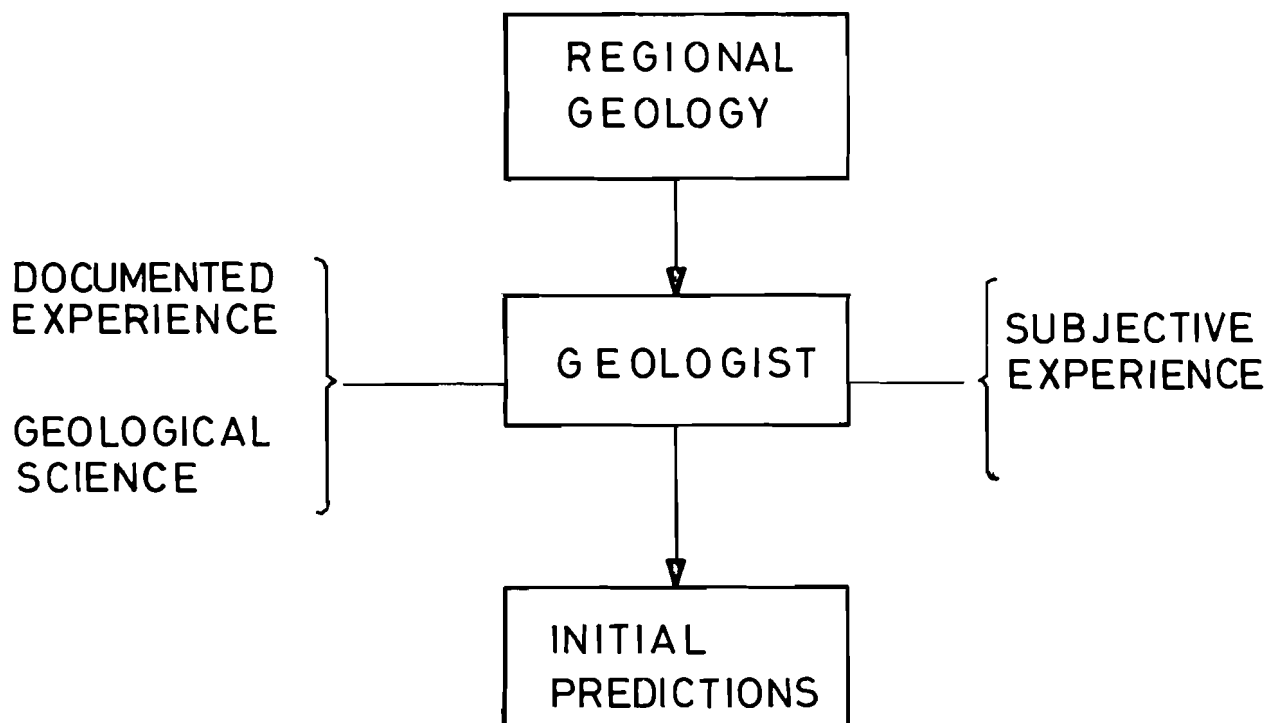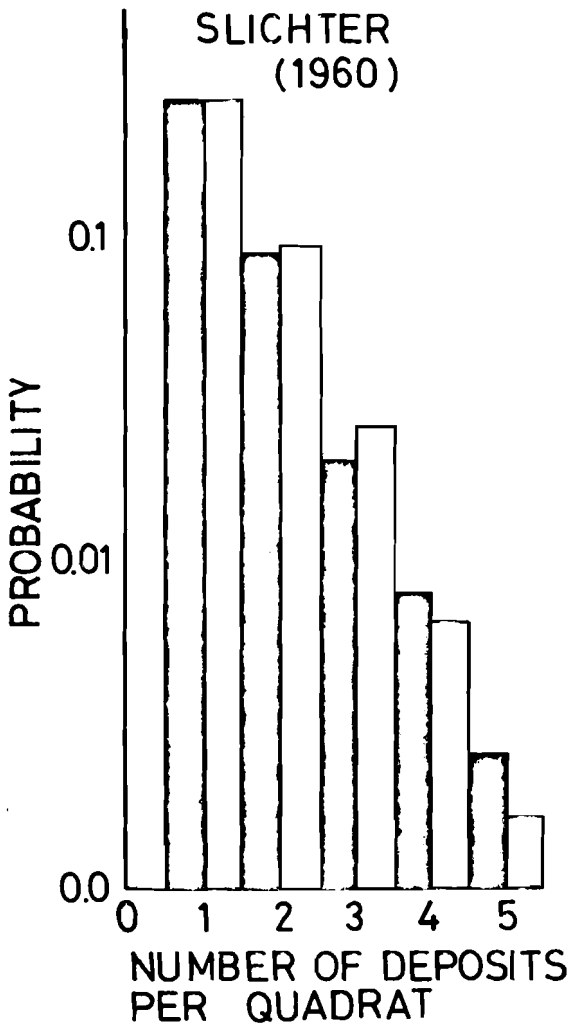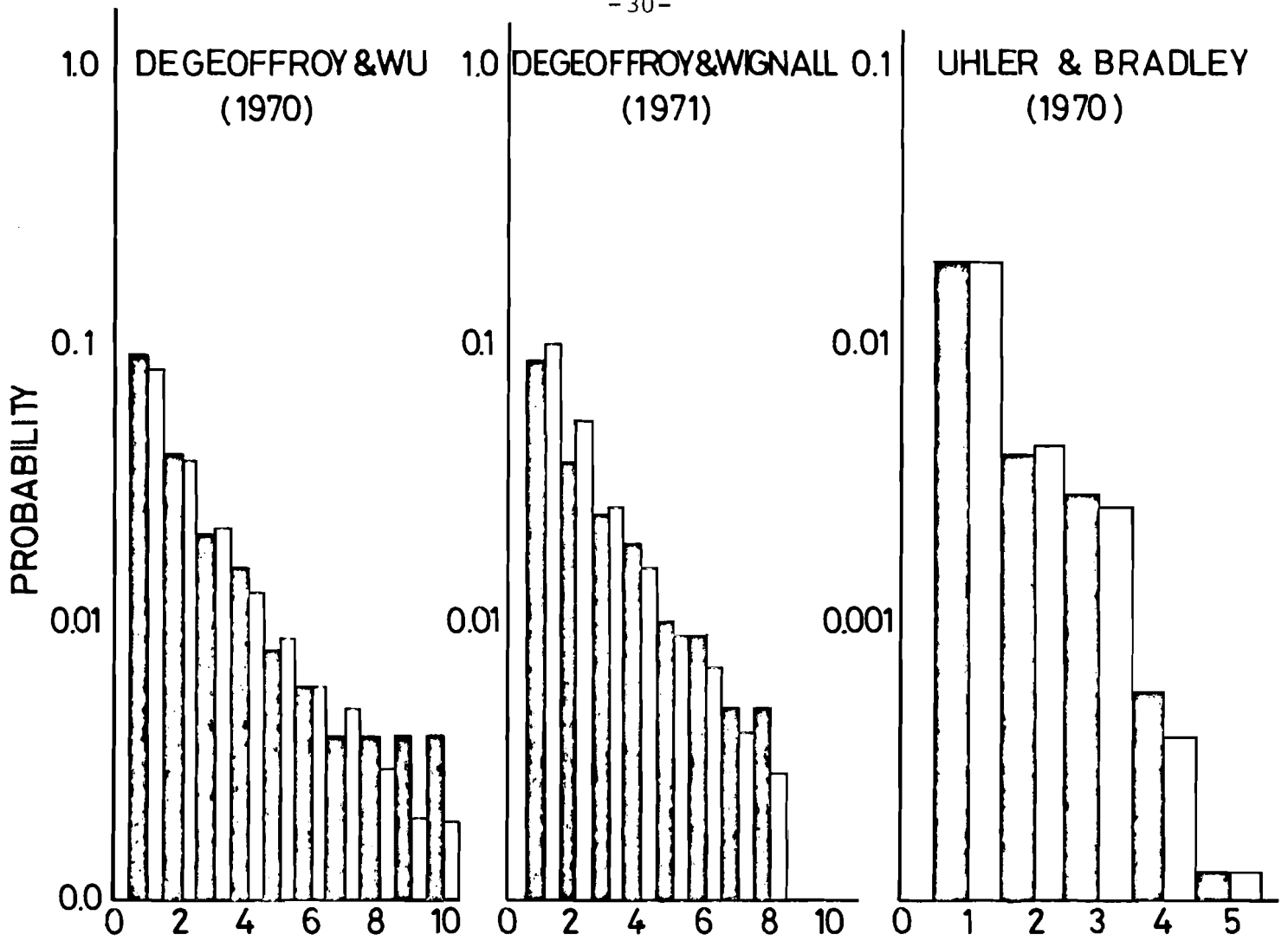
FIGURE 1

FIGURE 2

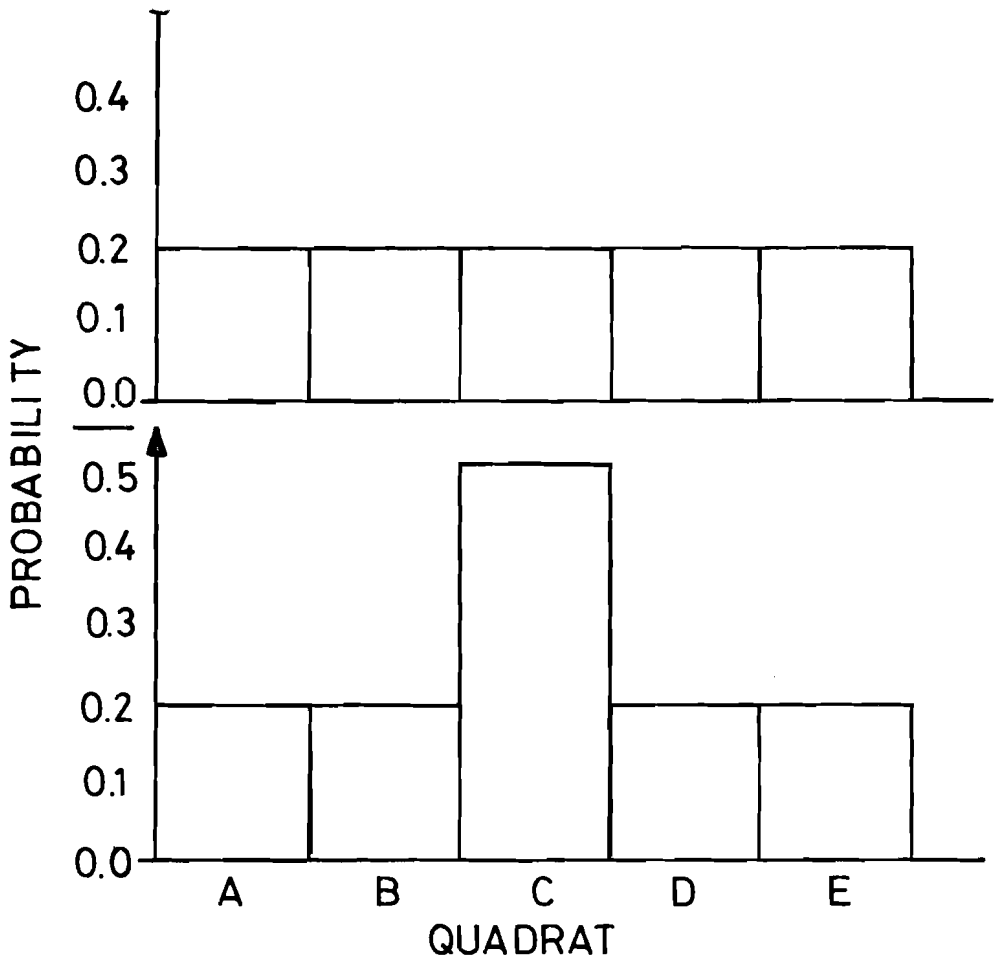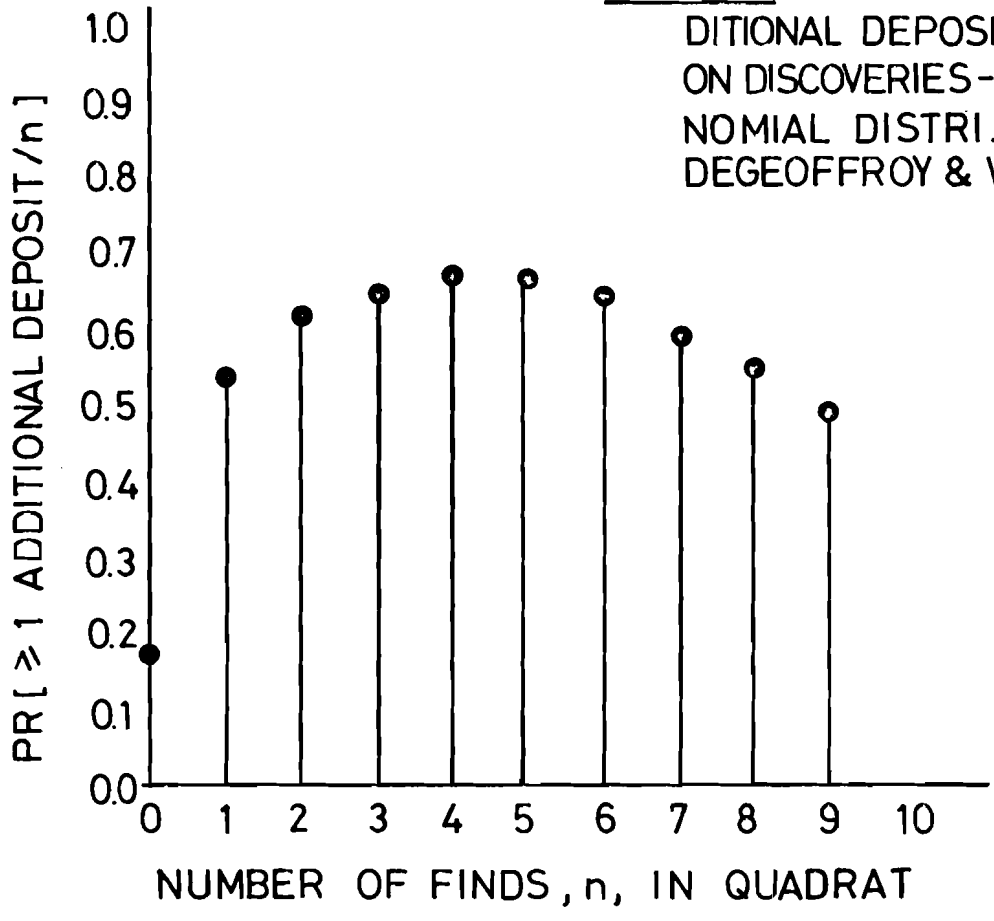FIGURE 3: FIT OF THE NEGATIVE BINOMINAL TO EMPIRICAL DATA

FIGURE 4:PROBABILITIES OF ≥1 AD-
DITIONAL DEPOSIT CONDITIONED
ON DISCOVERIES - NEGATIVE BI-
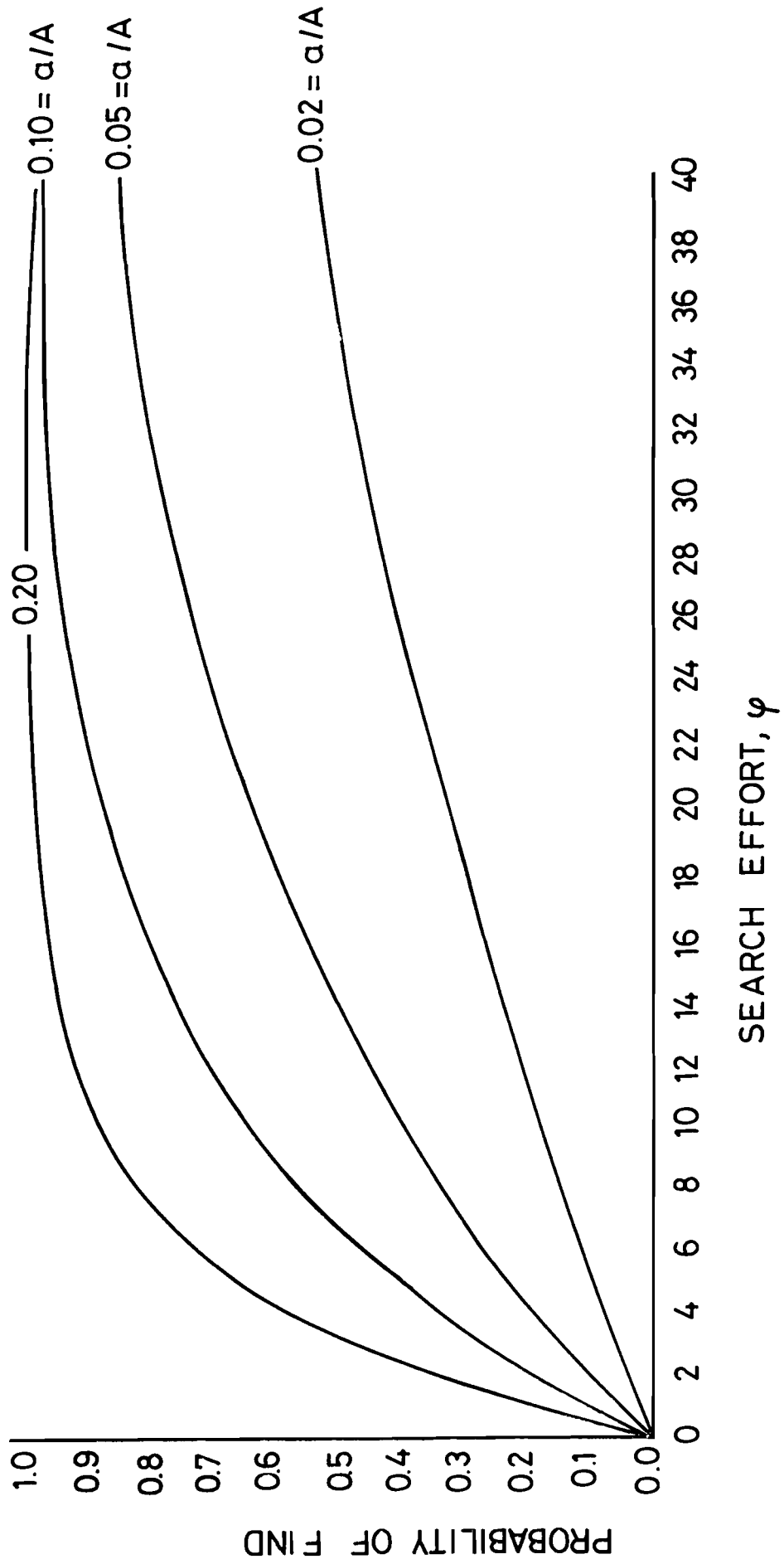NOMIAL DISTRI., DATA AFTER
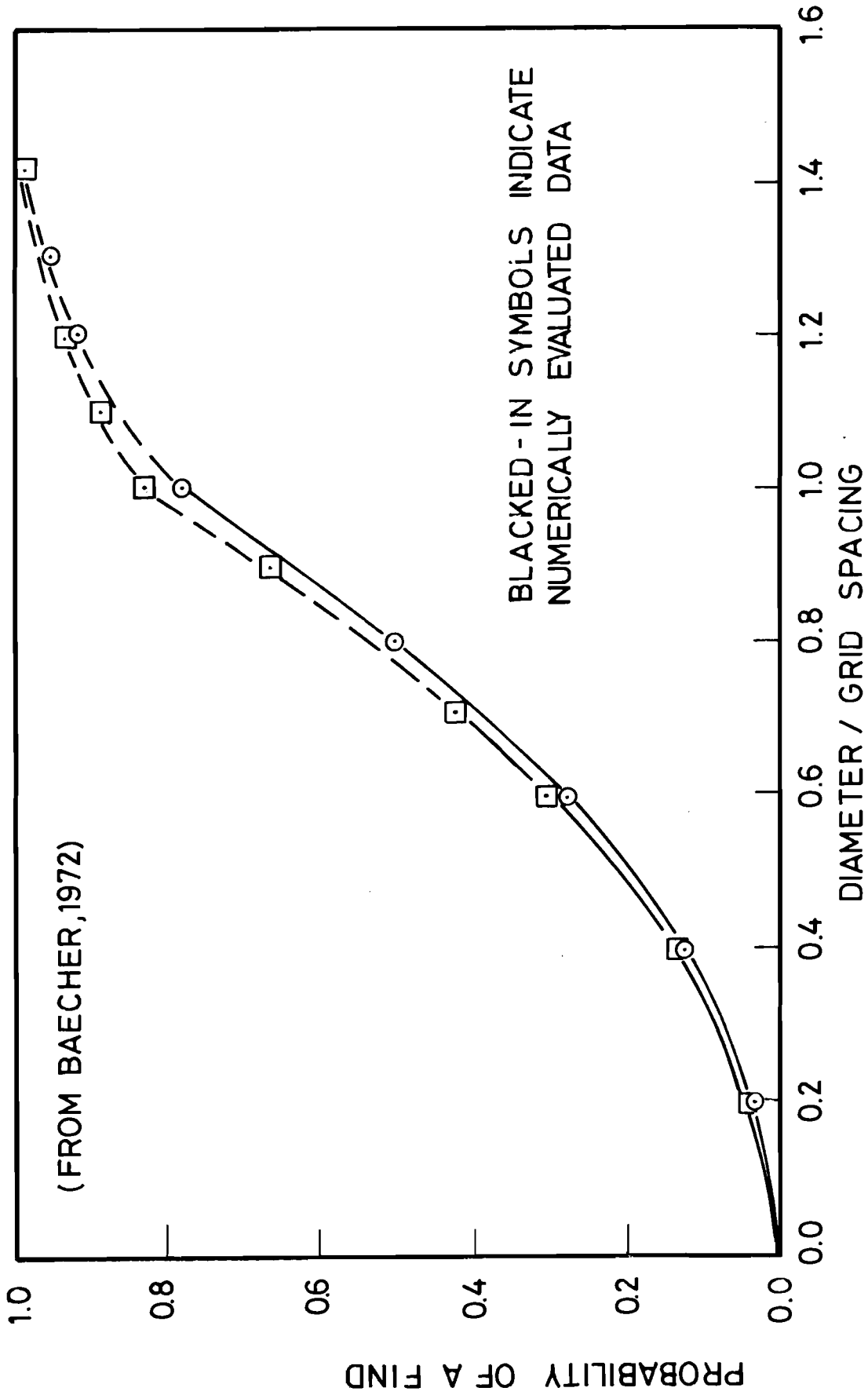DEGEOFFROY & WU ,1970

FIGURE 5 : RANDOM SEARCH

FIGURE 6a: PROBABILITY OF FINDING CIRCULAR AND SQUARE BODIES WITH SQUARE POINT GRID (EFFORT PROPORTIONAL TO (GRID SPACING)$^{-2}$)
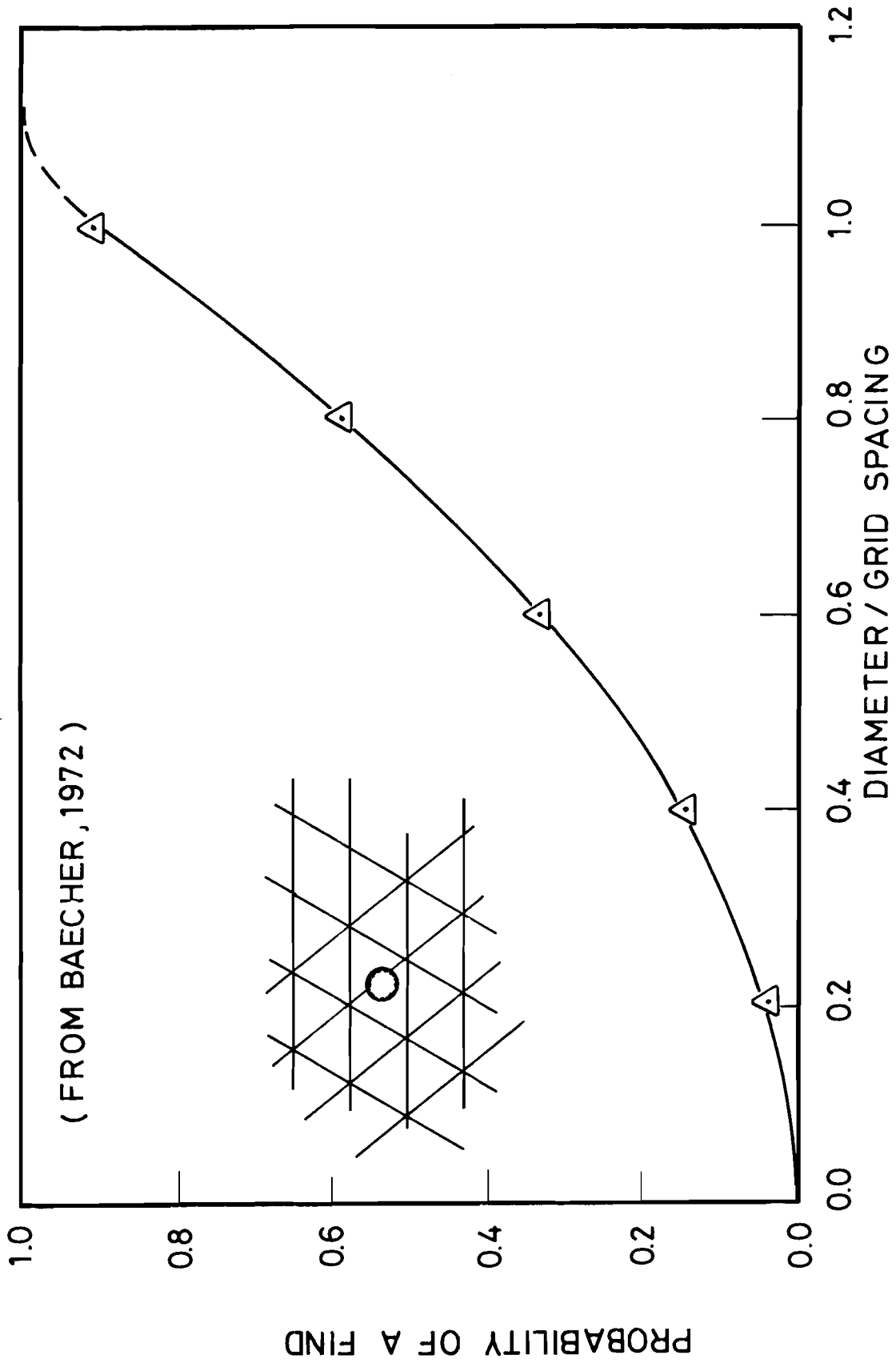
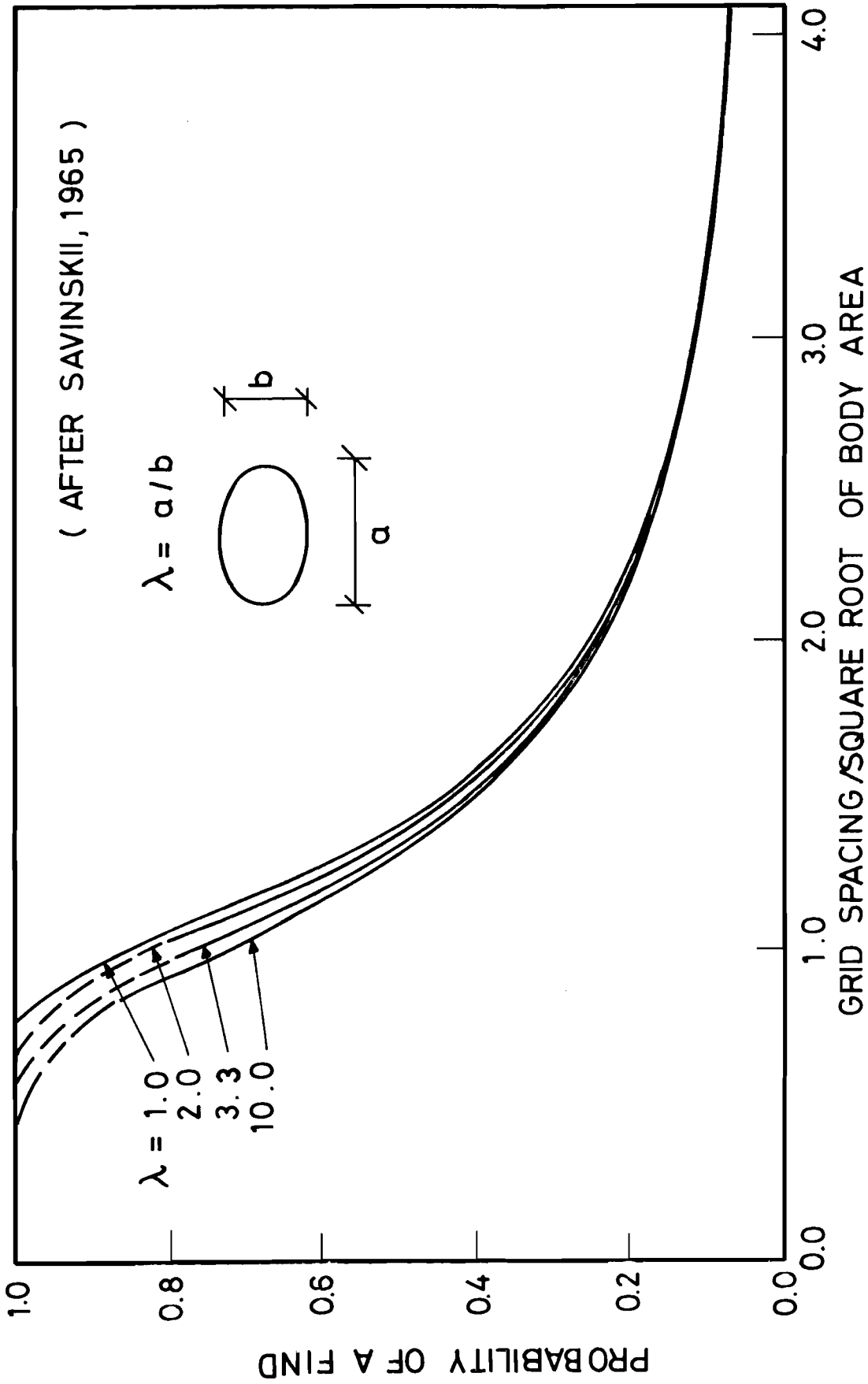FIGURE 6b: PROBABILITY OF FINDING BODY WITH TRIANGULAR POINT GRID.

FIGURE 6c: PROBABILITY OF FINDING ELLIPTICAL BODIES WITH SQUARE POINT GRIDS.

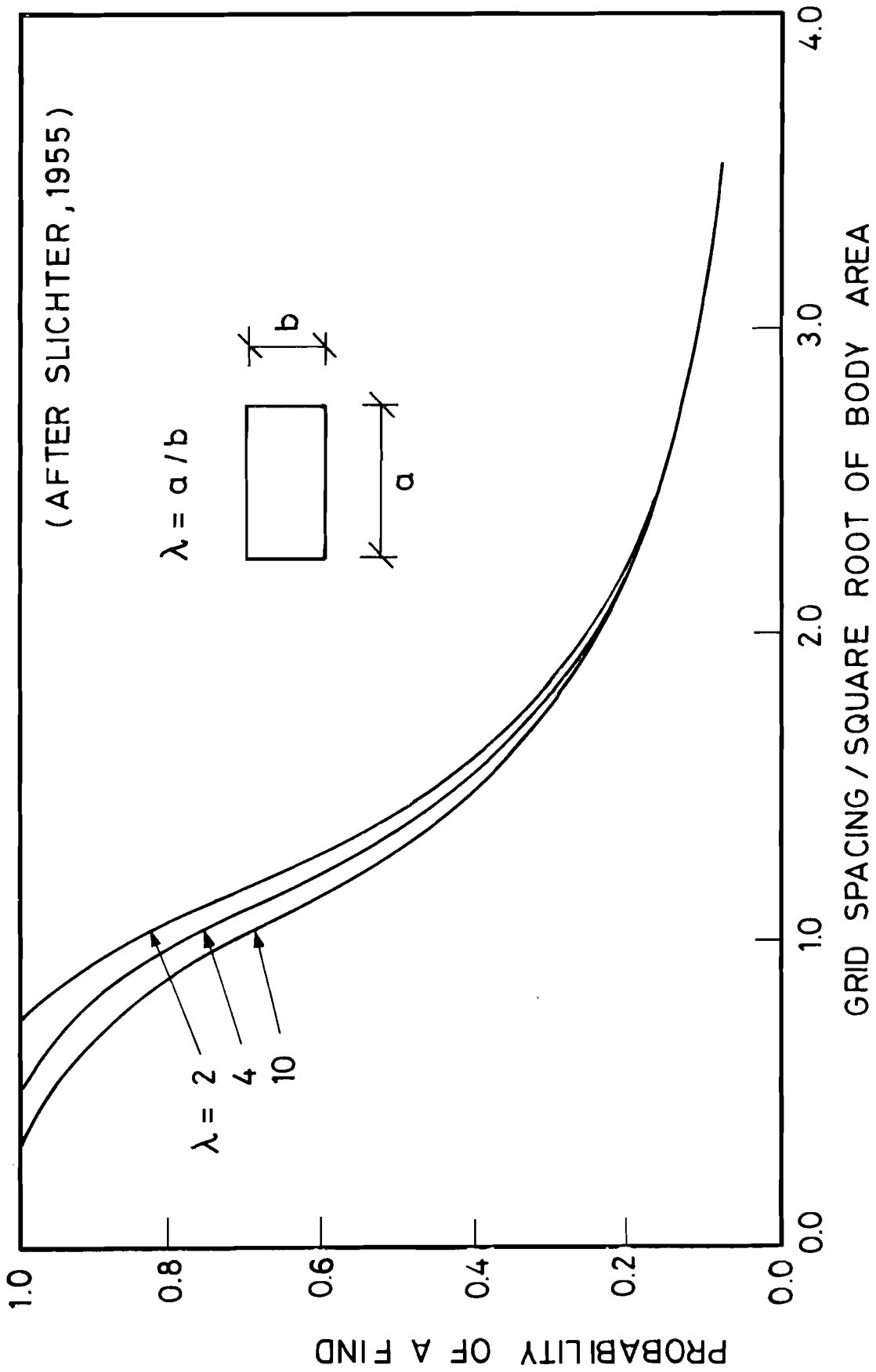FIGURE 6d: PROBABILITY OF FINDING RECTANGULAR BODY WITH SQUARE POINT GRID

References

[1] Allais, M. "Method of Appraising Economic Prospects
     of Mining Exploration over Large Territories."
     Management Science, 3 (1957), 285-347.

[2] Baecher, G.B. "Site Exploration: A probabilistic
     approach." Ph.D. dissertation, Massachusetts
     Institute of Technology, 1972.

[3] Box, G.E. and Tiao, G.C. Bayesian Inference in
     Statistical Analysis. Addison-Wesley, 1973.

[4] DeGeoffroy, J. and Wignall, T.K. "A Probabilistic
     Appraisal of Mineral Resources in a Portion of
     the Grenville Province of the Canadian Shield."
     Economic Geology, 66 (1971), 466-479.

[5] DeGeoffroy, J. and Wu, S.M. "A Statistical Study of
     Ore Occurrences in the Greenstone Belts of the
     Canadian Shield." Economic Geology, 65 (1970),
     496-504.

[6] Edwards, W. "Conservatism in Human Information
     Processing" in Formal Representation of Human
     Judgment. Ed., B. Kleinmontz. Wiley, 1968,
     pp. 17-52.

[7] Ericson, W.A. "Subjective Bayesian Models in Sampling
     Finite Populations." Jour. Roy. Stat. Soc., B, 31
     (1969), 195-233.

[8] Grayson, C.J. Decisions Under Uncertainty: Drilling
     Decisions by Oil and Gas Operators. Harvard
     University Business School, 1960.

[9] Griffiths, J.C. "Exploration for Natural Resources."
     Oper. Res., 14 (1966), 189-209.

[10] Griffiths, J.C. and Singer, D.A. "Unit Regional Value
     of Non-Renewable Natural Resources as a Measure of
     Potential for Development of Large Regions."
     Spec. Publs. Geol. Soc. Aust., 3 (1971), 227-238.

[11] Harris, D.P. "A Subjective Probability Appraisal of
     Metal Endowment of Northern Sonora, Mexico."
     Econ. Geol., 68 (1973), 222-242.

[12]   Harris, D.P. "An Application of Multivariate Statistical
        Analysis to Mineral Exploration." Ph.D. dissertation,
        Penn. State University, 1965.

[13]   Harris, D.P., Freyman, A.J., and Barry, G.S. "The
        Methodology Employed to Estimate Potential Mineral
        Supply of the Canadian Northwest--An Analysis Based
        upon Geologic Opinion and Systems Simulation."
        9th Intl. Symp. Techniques for Decision-Making in
        the Min. Ind., 1970.

[14]   Hubbert, M.K. "Energy Resources" in Resource and Man,
        National Academy of Sciences. Freeman and Company, 1969.

[15]   Kaufman, G.M. "Statistical Methods for Predicting the
        Number and Size of Undiscovered Hydrocarbon Deposits"
        in Methods of Estimating the Volume of Undiscovered
        Oil and Gas Resources, K.H. Crandall and J.W. Harbaugh.
        Stanford, 1974.

[16]   Kaufman, G.M. and Bradley, P.G. "Two Stochastic Models
        Useful in Petroleum Exploration." Am. Assn. Pet.
        Geol., Arctic Geol. Mem., 19 (1973), 633-637.

[17]   MacAvoy, P.W. and Pindyck, R.S. "Alternative Regulatory
        Policies for Dealing with the Natural Gas Shortage."
        Bell Jour. Mgt. Sci., 5 (1974), 454-498.

[18]   Moore, C.L. "Limitations of Statistical Methods for
        Predicting Petroleum and Natural Gas Reserves and
        Availability--A Reply." Jour. Pet. Technology, 18, 3
        (1966), 286.

[19]   Morris, P.A. Bayesian Expert Resolution. Ph.D. dissertation,
        Stanford University, 1971.

[20]   Morris, P.A. "Decision Analysis Expert Use." Management
        Science, 20 (1974), 1233-1241.

[21]   Morse, P.M. Search Theory. Oper. Res. Center, MIT.
        Working Paper 01-74, 1974.

[22]   Murphy, A.H. and Winkler, R.L. "Credible Interval Temperature
        Forecasting: Some Experimental Results." Monthly
        Weather Review, 102 (1974), 784-794.

[23]   Pill, J. "The Delphi Method: Substance, Context,
        A Critique, and An Annotated Bibliography." Socio.-
        Econ. Plan. Sci., 5 (1971), 57-71.

[24]   Robinson, J.M.   "Nature and Significance of Geological
            Maps" in Fabric of Geology. Ed., C.C. Albritton.
            Freeman and Company, 1963.

[25]   Rogers, A.   Spatial Dispersion.   Pion Monographs in
            Environmental and Spatial Systems.   Pion, 1974.

[26]   Ryan, J.T.   "An Analysis of Crude-Oil Discovery Rate
            in Alberta."   Bull. Can. Pet. Geol., 21 (1973),
            219-235.

[27]   Ryan, J.T.   "An Estimate of the Conventional Crude-Oil
            Potential in Alberta."   Bull. Can. Pet. Geol., 21
            (1973), 236-246.

[28]   Savinskii, J.   Probability Tables for Locating Elliptical
            Underground Masses with a Rectangular Grid.   New York,
            Consultants Bureau, 1965.

[29]   Singer, D.A. and Wickman, F.E.   "Probability Tables for
            Locating Elliptical Targets with Square, Rectangular
            and Hexagonal Point-Nets."   Mineral Sci. Exp. Sta.,
            Penn. State University.   Spec. Pub. 1-69, 1969.

[30]   Slichter, L.B.   "The Need of a New Philosophy of Prospecting."
            Min. Engr., 12 (1960), 570-576.

[31]   Tversky, A.   "Assessing Uncertainty."   Jour. Roy. Stat. Soc.,
            36 (1974), 148-160.

[32]   Uhler, R.S. and Bradley, P.G.   "A Stochastic Model for
            Determining the Economic Prospects of Petroleum
            Exploration over Large Regions."   Jour. Am. Stat. Assn.,
            65 (1970), 623-630.

[33]   Winkler, R.L. and Murphy, A.H.   "On the Generalizability
            of Experimental Results" in The Concept of Probability
            in Psychological Experiments. Ed., C.-A.S. Staël von
            Holstein.   Dordrecht-Holland, Reidel Publishing Company,
            1974, pp. 103-126.

[34]   Wood, E.F.   "A Bayesian Approach to Analyzing Uncertainty
            Among Stochastic Models."   International Institute for
            Applied Systems Analysis, Laxenburg, Austria, RR-74-16,
            1974.