

STATISTICAL MODELS FOR TOPOLOGICAL ANALYSIS
OF CUT DENDRITIC TREES
by Myra L. Samuels

Department of Statistics
Division of Mathematical Sciences
Purdue University
Mimeograph Series No. 512
December 1977

CONTENTS

	Page
1. Introduction	
1.1 Purpose	1
1.2 The Problem	1
1.3 Outline of the Report	2
2. First-Order Branching	4
3. Second-Order Branching	
3.1 Formulation of the MBC λ Model	7
3.2 The Partial Non-Identifiability of λ_2	13
3.3 Maximum Likelihood Estimation of All Parameters	16
3.4 Estimation and Testing for Fixed λ_2	19
4. Extension of the Models to All Orders of Branching	23
5. Analysis of Data	
5.1 Numerical Results	29
5.2 Summary of Data Analysis	48
Appendix	50
References	52

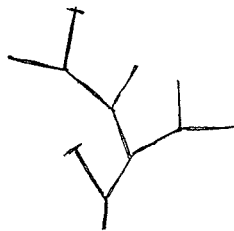
1. INTRODUCTION

1.1. Purpose. It is the purpose of this report to detail the methods used to produce the parameter estimates reported in "Complexity of Branching in Dendritic Trees: Dependence on Number of Trees per Cell and Effects of Branch Loss During Sectioning" [2]. This report is intended for biologically oriented readers of [2]. A briefer description of the statistical methods is currently in preparation for submission to Biometrics.

1.2. The Problem. The problem addressed is the estimation of the topological properties (branching structure) of a population of trees, some of whose branches are cut. The model is designed to fit the problem of the cutting of dendritic branches of neurons by the microtome in the process of preparing the histological section from which the data is obtained.

The following data structure is assumed: There is a population of trees, homogeneous in the sense that the probabilistic rules governing tree branching structure are the same for each tree in the population.¹ From this population a random sample of T trees is drawn. Some of the branches of these trees are cut, so that the part of the tree structure lying beyond each cut is missing.

The data consists of T drawings, each showing the branching structure of the uncut part of a tree, and also showing which branches are cut. Here is a typical (schematic) drawing:



The slashes indicate cut branches.

¹In the application to dendrites, each "tree" is the dendritic arborization of a single primary dendrite. It is shown in [2] that the dendritic branching probabilities depend considerably upon the number of trees per neuron. Consequently, in order to satisfy the requirement of homogeneity of probabilistic rules, it is necessary to group the data according to the number of trees per neuron, and analyze each group separately. In the data used in [2], the groups " ≤ 4 trees/cell", "5 trees/cell", and " > 6 trees/cell" were used. This grouping reduces, but does not eliminate, heterogeneity; further breakdown of the first and third categories was not feasible because of the small sample sizes involved.

In this report, each branch-segment between bifurcation nodes will be termed a "branch". The branches of a tree are ordered as follows: the primary dendrite (the "trunk" of the tree) is the first-order branch; the two (if any) arising from it are second-order branches, and so forth. Branching and cutting probabilities are assumed to be dependent on the order of the branch.

1.3. Outline of the Report. In [2], populations of trees are described in terms of parameters N_k , defined as:

$$N_k = \text{expected (population mean) number of } k^{\text{th}}\text{-order branches per tree.}$$

Of course, $N_1 = 1$. The primary goal of the methods of the present report is the estimation, from cut tree data, of N_2, N_3, \dots , under a reasonable model.

The problem of statistical inference from cut tree data is discussed separately for the various orders of branching. In Section 2 estimation of first-order branching and cutting parameters (including N_2) is discussed. It is found that this estimation requires an arbitrary parameter, λ_1 , which cannot be estimated from the data.

Section 3 is devoted to several models for estimation of second-order branching and cutting parameters, including N_3 . These models involve a parameter λ_2 , which has a paradoxical property: under a certain submodel, the BB & BC model, the parameter λ_2 is non-identifiable, i.e., it cannot be estimated from topological data; under a larger model, the MBC λ model, λ_2 can be estimated (but only very poorly if the data fits the BB & BC submodel at all well). In Section 3.1 the MBC λ model is formulated. In Section 3.2 the paradoxical "partial non-identifiability" of λ_2 is discussed. Section 3.3 deals with the estimation of all second-order branching and cutting parameters,

including λ_2 . Section 3.4 details procedures for estimating second-order branching and cutting parameters when λ_2 is arbitrarily fixed, rather than estimated from data. The N_3 estimates reported in [2] were calculated by the methods of Section 3.4, since the data available to us allowed only very poor estimation of λ_2 . Also in Section 3.4, procedures are described for testing the goodness of fit of data to the BB & BC model, the MBC λ model, and to two intermediate models.

In Section 4 one of these intermediate models, the BB model, is extended to cover all orders of branching. Section 4 culminates in equations for calculating estimates of N_k , for any k , under the extended BB model. Such estimates depend upon arbitrarily prespecified parameters λ_k .

In Section 5.1, the data collected by Professor P. D. Coleman and used in [2] is summarized, and some of the results of the analysis of this data are presented. These results include: goodness-of-fit statistics showing that λ_2 cannot be well estimated from the data (Table IV); goodness-of-fit statistics showing that the second-order branching data fits the BB model fairly well (Table V) and a comparison table showing that the estimates of N_3 under the BB model are quite close to those under the full MBC λ model (Table VI); goodness-of-fit statistics showing that (as would be expected) the other intermediate model, the BC model, is not well-fit by the second-order data (Table VII); values of the estimates of N_3 for selected values of λ_1 and λ_2 (Table VIII); values of the estimates of N_2, N_3, N_4 , and N_5 , and also of W_2, W_3, W_4 , and W_5 (where W_k = mean number of k^{th} -order branches per neuron) for $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0, .5, 1, 2, 4$ and ∞ . (Tables IX - XIV). In Section 5.2 the main results of the data analysis are briefly summarized.

Let β_1 be the bifurcation probability for a first-order branch. Also, let

$$t_1 = P\{\text{a first-order terminal branch is cut}\}$$

$$b_1 = P\{\text{a first-order bifurcating branch is cut}\}.$$

The probability of cutting of a branch is allowed to depend upon the bifurcation or non-bifurcation of the branch because of the evidence in the literature (discussed in [2]) that terminal branches tend to be longer than bifurcating branches. Longer branches presumably have a greater chance of being cut.

The three distinguishable configurations of the first two orders of branches (after cutting) are shown in Figure 1,² together with a notation for the probabilities of occurrence and the observed frequencies of occurrence.

Note that $x_1 + y_1 + z_1 = T =$ the total number of trees in the sample.


Configuration		/	
Probability	ϕ_1	ψ_1	$1 - \phi_1 - \psi_1$
Observed Frequency	x_1	y_1	z_1

Figure 1

It is easy to see that

$$\phi_1 = \beta_1(1 - b_1)$$

and
$$\psi_1 = (1 - \beta_1)(1 - t_1). \quad (2.1)$$

The probabilities ϕ_1 and ψ_1 can be estimated from x_1 , y_1 , and z_1 ; however, the parameters β_1 , b_1 , and t_1 cannot all be estimated because the values of ϕ_1 and ψ_1 do not uniquely determine β_1 , b_1 and t_1 . In fact, it can be seen from Equation (2.1) that any given pair (ϕ_1, ψ_1) is equally compatible with the hypothesis $t_1 = 0$ (no terminal branches are cut) and the hypothesis $b_1 = 0$ (no bifurcating branches are cut). Of course, these two hypotheses imply very different values of β_1 .

²In Figure 1, and in subsequent figures in this report, a slash indicates a cut branch; a dot at the end of a branch indicates that the branch may be cut or uncut, and may or may not bifurcate. Unmarked branches are terminal.

In order to estimate β_1 , therefore, it is necessary to make some assumption about t_1 and b_1 . Let

$$\lambda_1 = t_1/b_1. \quad (2.2)$$

Then, as we have seen, any value of λ_1 , $0 \leq \lambda_1 \leq \infty$, is equally compatible with a given set of data. If, however, a value of λ_1 is specified arbitrarily, then β_1 , b_1 and t_1 can be estimated from the data. If we estimate ϕ_1 and ψ_1 by

$$\begin{aligned} \hat{\phi}_1 &= x_1/T \\ \hat{\psi}_1 &= y_1/T, \end{aligned} \quad (2.3)$$

(which are the maximum likelihood estimates), then the estimates $\hat{\beta}_1$, \hat{b}_1 , and \hat{t}_1 of β_1 , b_1 and t_1 are determined from the equations

$$\hat{\beta}_1^2(1-\lambda_1)T - \hat{\beta}_1[x_1 + z_1 - \lambda_1(x_1+T)] - \lambda_1x_1 = 0 \quad (2.4)$$

$$\begin{aligned} \hat{b}_1 &= 1 - x_1/\hat{\beta}_1T \\ \hat{t}_1 &= \lambda_1\hat{b}_1. \end{aligned} \quad (2.5)$$

Equations (2.4) and (2.5) are easily obtained from Equations (2.1), (2.2) and (2.3). It is easy to show that, if x_1 and y_1 are both nonzero, then for each value of λ_1 , $0 \leq \lambda_1 < \infty$, (2.4) has a unique solution $\hat{\beta}_1$ between 0 and 1. In the extreme case $\lambda_1 = \infty$ (i.e., $b_1 = 0$), Equations (2.4) and (2.5) must be replaced by

$$\hat{\beta}_1 = \frac{x_1}{T} \quad (2.4')$$

$$\hat{t}_1 = \frac{z_1}{y_1+z_1} \quad (2.5')$$

In [2] we consider the quantity N_2 , defined as the expected (population mean) number of the second-order branches per tree. In terms of the present notation, $N_2 = 2\beta_1$, so that N_2 may be estimated by

$$\hat{N}_2 = 2\hat{\beta}_1. \quad (2.6)$$

In summary, we have seen in this section that the estimation of β_1 , the branching probability for first-order branches, requires the arbitrary parameter λ_1 , whose value must be guessed. If terminal branches have the same chance of being cut as bifurcating branches, then $\lambda_1 = 1$. If terminal branches are longer and hence more prone to cutting, then $\lambda_1 > 1$. We shall see in the next section that the estimation of second-order branching probabilities may require an arbitrary parameter λ_2 , or it may not, depending upon the nature of the data set.

3. SECOND-ORDER BRANCHING

3.1. Formulation of the MBC λ Model

In order to estimate the frequency with which second-order branches bifurcate to produce third-order branches, we first avoid consideration of first-order branches by conditioning on the event E, defined as follows³:

$$E = \{\text{first-order branch bifurcates and is not cut}\}.$$

All probabilities and stochastic assumptions in this section are conditioned on the event E; for the sake of brevity, this conditioning will not be explicitly mentioned in the definitions.

Given that the first-order branch bifurcates, we can define branching probabilities for the two second-order branches that arise from it. Let⁴

$$p_{11} = P\{\text{neither second-order branch bifurcates}\}$$

$$2p_{12} = P\{\text{exactly one second-order branch bifurcates}\}$$

$$p_{22} = P\{\text{both second-order branches bifurcate}\}.$$

Note that we do not make the assumption that the two second-order branches bifurcate independently of each other. It turns out that this assumption can be tested from data; therefore we state it here as a hypothesis for future reference:

Binomial Branching Hypothesis H_{BB} : The two second-order branches bifurcate independently of each other; i.e., there is a second-order "branching

³Since all the estimates considered in this report are maximum likelihood estimates, this conditioning poses no difficulties. The unconditional likelihood function factors suitably, so that the estimation of the conditional second-order parameters by maximization of the conditional likelihood function gives estimates which are independent of the first-order parameter estimates, and the estimates so obtained maximize the unconditional likelihood function.

⁴The meanings of the p_{ij} are different from their meanings in [1]. In [1], the p_{ij} were not conditional on E.

probability" β_2 , such that

$$\begin{aligned} p_{11} &= (1 - \beta_2)^2 \\ p_{12} &= \beta_2(1 - \beta_2) \\ p_{22} &= \beta_2^2 \end{aligned} \quad (2.1)$$

Next we consider the cutting of the two second-order branches. It seems likely that cutting of a second-order branch is not independent of the cutting of its sister second-order branch; rather, since there is only one cutting plane (the plane of section)⁵, there is positive correlation between the two branches with respect to cutting. Motivated by this consideration, we have constructed the Modified Binomial Cutting (MBC) model, which takes the cutting plane into account.

The basic assumption of the MBC model is that there is a vector random variable R such that, conditional on the value r of R , the two sister branches are cut or not cut independently of each other. The motivating idea is that r is some description of the position of the entire dendritic tree with respect to the plane of section. For example, r could be the vector originating at the node where the first-order branch bifurcates and orthogonal to the cutting plane, as indicated in Figure 2.

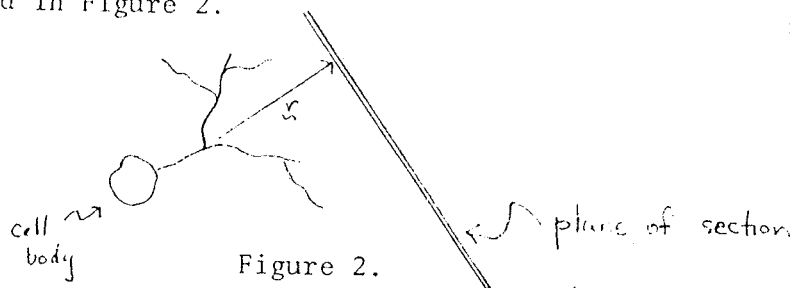


Figure 2.

⁵The chance of the two parallel planes of section both cutting the same dendritic tree is considered to be vanishingly small.

The strength of the MBC model, however, is that \underline{r} need not be defined precisely. It is sufficient to assume the existence of the following functions:

$$t_2(\underline{r}) = P\{\text{a second-order terminal branch is cut} | \underline{r}\}$$

$$b_2(\underline{r}) = P\{\text{a second-order bifurcating branch is cut} | \underline{r}\}$$

$$F(\underline{r}) = \text{probability distribution function}^6 \text{ of } \underline{R} = P\{\underline{R} \leq \underline{r}\}$$

The forms of the functions $t_2(\cdot)$, $b_2(\cdot)$, and $F(\cdot)$ are determined by such factors as the type of neuron, the method of preparing the section, and (probably most important) the method of selecting a particular dendritic tree to be included in the data.

We shall find it relevant later to consider the hypothesis that two sister second-order branches are cut independently of each other. This can be realized as a special case of the MBC model by assuming that $t(\underline{r}) \equiv t_2$ and $b(\underline{r}) \equiv b_2$. For future reference, we state here this hypothesis and the hypothesis obtained by conjoining it with the Binomial Branching Hypothesis.

Binomial Cutting Hypothesis H_{BC} : The two second-order branches are cut or not cut independently of each other, with cutting probabilities t_2 for terminal branches and b_2 for bifurcating branches.

Binomial Branching and Cutting Hypothesis $H_{BB\&BC}$: both H_{BB} and H_{BC} hold.

Armed with the MBC assumptions, we can consider the probabilities of various transitions from third-order branch configurations before cutting to configurations after cutting. The following transition probability

⁶The vector inequality $\underline{R} \leq \underline{r}$ means coordinate-wise inequality.

matrix summarizes these configurations⁷; the interior of the matrix is left blank except for zeros indicating impossible transitions.

Distinguishable Configurations after Cutting



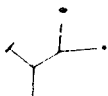


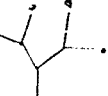


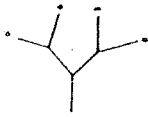
Configurations before Cutting	     																		
  	<table border="0" style="margin: auto;"> <tr> <td style="text-align: center;">-</td> <td style="text-align: center;">-</td> <td style="text-align: center;">0</td> <td style="text-align: center;">-</td> <td style="text-align: center;">0</td> <td style="text-align: center;">0</td> </tr> <tr> <td style="text-align: center;">-</td> <td style="text-align: center;">-</td> <td style="text-align: center;">-</td> <td style="text-align: center;">0</td> <td style="text-align: center;">-</td> <td style="text-align: center;">0</td> </tr> <tr> <td style="text-align: center;">-</td> <td style="text-align: center;">0</td> <td style="text-align: center;">-</td> <td style="text-align: center;">0</td> <td style="text-align: center;">0</td> <td style="text-align: center;">-</td> </tr> </table>	-	-	0	-	0	0	-	-	-	0	-	0	-	0	-	0	0	-
-	-	0	-	0	0														
-	-	-	0	-	0														
-	0	-	0	0	-														

Figure 3. Transition Matrix

The missing transition probabilities in Figure 3 can all be easily expressed in terms of the MBC model. For example, the transition



has probability

$$\int t_2(\mathbf{r}) [1 - b_2(\mathbf{r})] dF(\mathbf{r})$$

$$= \bar{t}_2 - \overline{t_2 b_2},$$

where the bar "-" indicates integration with respect to the distribution $F(\mathbf{r})$. In a similar way, all the transition probabilities can be expressed in terms of the 5 parameters

$$\bar{t}_2 = \int t_2(\mathbf{r}) dF(\mathbf{r})$$

$$\bar{b}_2 = \int b_2(\mathbf{r}) dF(\mathbf{r})$$

⁷ See Footnote 2 for explanation of tree drawings.

$$\bar{t}_2^2 = \int [t_2(\underline{r})]^2 dF(\underline{r}) \quad (3.2)$$

$$\bar{b}_2^2 = \int [b_2(\underline{r})]^2 dF(\underline{r})$$

and
$$\overline{t_2 b_2} = \int t_2(\underline{r}) b_2(\underline{r}) dF(\underline{r}).$$

The estimation problem for second-order branching can now be explicitly presented. Figure 4 shows the 6 distinguishable configurations of third-order branches after cutting, together with a notation for the observed frequencies and the probabilities.







Configuration						
Probability	κ	ν_1	ν_2	μ_{11}	μ_{12}	μ_{22}
Observed Frequency	k	n_1	n_2	m_{11}	m_{12}	m_{22}

Figure 4.

We shall refer to the probabilities of the observable tree categories as the $\{\kappa\nu\mu\}$ parameters. Referring to Figures 3 and 4, it is easy to see that the $\{\kappa\nu\mu\}$ parameters can be written in terms of the parameters of the MBC model as⁸:

⁸The five parameters \bar{t}_2 , \bar{t}_2^2 , \bar{b}_2 , \bar{b}_2^2 , and $\overline{t_2 b_2}$ are subject to two sets of constraints. The first set, namely,

$$0 \leq \bar{t}_2 - \bar{t}_2^2 \leq 1 - \bar{t}_2, \quad 0 \leq \bar{b}_2 - \bar{b}_2^2 \leq 1 - \bar{b}_2, \quad 0 \leq \bar{t}_2 - \overline{t_2 b_2} \leq 1 - \bar{b}_2,$$

$$0 \leq \bar{b}_2 - \overline{t_2 b_2} \leq 1 - \bar{t}_2,$$

suffice to guarantee that the transition probabilities (the coefficients of the p_{ij} in (3.3)) are nonnegative. A second set of constraints, namely,

$$\bar{t}_2^2 \geq (\bar{t}_2)^2, \quad \bar{b}_2^2 \geq (\bar{b}_2)^2, \quad \overline{t_2 b_2} \geq (\bar{t}_2)(\bar{b}_2), \quad \bar{t}_2 - \bar{t}_2^2 \leq 1/4, \quad \bar{b}_2 - \bar{b}_2^2 \leq 1/4,$$

is implied by the interpretation of the five parameters in terms of $t_2(\underline{r})$, $b_2(\underline{r})$, and $F(\underline{r})$. Since the probabilities (3.3) are meaningful without the second set of constraints, we will broaden the model by dropping the second set of constraints.

$$\begin{aligned}
\kappa &= p_{11}\overline{t_2^2} + 2p_{12}\overline{t_2 b_2} + p_{22}\overline{b_2^2} \\
\nu_1 &= 2[p_{11}(\overline{t_2} - \overline{t_2^2}) + p_{12}(\overline{b_2} - \overline{t_2 b_2})] \\
\nu_2 &= 2[p_{12}(\overline{t_2} - \overline{t_2 b_2}) + p_{22}(\overline{b_2} - \overline{b_2^2})] \\
\mu_{11} &= p_{11}(1 - 2\overline{t_2} + \overline{t_2^2}) \\
\mu_{12} &= 2p_{12}(1 - \overline{t_2} - \overline{b_2} + \overline{t_2 b_2}) \\
\mu_{22} &= p_{22}(1 - 2\overline{b_2} + \overline{b_2^2}).
\end{aligned} \tag{3.3}$$

The statistical problem, then, is estimation of the parameters of the model from the observed frequencies k , n_1 , n_2 , m_{11} , m_{12} , and m_{22} . However, the model as it stands contains too many parameters to be estimated from these six observed frequencies. It is necessary to make some additional assumption.

The additional assumption we shall make is the Constant-Ratio Assumption: There is a constant, λ_2 , such that, for all r ,

$$\frac{t_2(r)}{b_2(r)} = \lambda_2. \tag{3.4}$$

Note that the Constant-Ratio Assumption is unrealistic, since presumably both $t_2(r)$ and $b_2(r)$ approach unity as $r \rightarrow 0$. It is to be hoped that this difficulty has relatively little effect on the $\{\kappa\nu\}$ parameters. The author intends in a future report to consider another version of the MBC model with (3.4) replaced by another assumption. In the present report, the MBC model incorporating (3.4) will be used; for brevity, we shall call this the MBC λ model.

The Constant-Ratio Assumption implies that

$$\begin{aligned} \overline{t_2} &= \lambda_2 \overline{b_2} \\ \overline{t_2} &= \lambda_2^2 \overline{b_2^2} \\ \overline{t_2 b_2} &= \lambda_2 \overline{b_2^2} \end{aligned} \quad (3.5)$$

From (3.3), then, the $\{\kappa\nu\mu\}$ parameters can be written under the MBC λ model as:

$$\begin{aligned} \kappa &= \overline{b_2^2} (\lambda_2^2 p_{11} + 2 \lambda_2 p_{12} + p_{22}) \\ \nu_1 &= 2 (\lambda_2 p_{11} + p_{12}) (\overline{b_2} - \lambda_2 \overline{b_2^2}) \\ \nu_2 &= 2 (\lambda_2 p_{12} + p_{22}) (\overline{b_2} - \overline{b_2^2}) \\ \mu_{11} &= p_{11} (1 - 2 \lambda_2 \overline{b_2} + \lambda_2^2 \overline{b_2^2}) \\ \mu_{12} &= 2 p_{12} (1 - \lambda_2 \overline{b_2} - \overline{b_2} + \lambda_2 \overline{b_2^2}) \\ \mu_{22} &= p_{22} (1 - 2 \overline{b_2} + \overline{b_2^2}) \end{aligned} \quad (3.6)$$

3.2. The Partial Non-Identifiability of λ_2

In the MBC λ model, the number of parameters appears to be under control. There are 5 independent parameters in the model: λ_2 , $\overline{b_2}$, $\overline{b_2^2}$, p_{12} , and p_{22} . (recall that $p_{11} = 1 - 2p_{12} - p_{22}$). The distribution of the observed frequencies (k , n_1 , n_2 , m_{11} , m_{12} , m_{22}) is 6-nomial; i.e., multinomial with 5 degrees of freedom. It would appear, then, that all the parameters can be estimated. Unfortunately, the situation is not that simple. A problem of identifiability, similar to that observed in Section 2, still lurks in the shadows.

In Section 2, the parameter λ_1 was seen to be non-identifiable; that is, given values of ϕ_1 and ψ_1 do not determine a unique value of λ_1 , so that estimation of ϕ_1 and ψ_1 does not provide an estimate of λ_1 . For second-order branching, the identifiability situation is more complicated. We shall see that, while some configurations of the $\{\kappa\nu\mu\}$ probabilities are compatible with any value of λ_2 from 0 to ∞ , other configurations restrict, or even uniquely determine, the value of λ_2 .

In order to describe the identifiability situation for second-order branching, let us consider the submodel specified by the hypothesis $H_{BB\&BC}$ stated earlier in this section. $H_{BB\&BC}$ implies the Constant-Ratio Assumption and also that

$$\begin{aligned}\overline{b_2} &= b_2 \\ \overline{b_2^2} &= b_2^2\end{aligned}\tag{3.7}$$

Using (3.6) and (3.7), it follows readily that under the hypothesis $H_{BB\&BC}$ all of the $\{\kappa\nu\mu\}$ probabilities (3.3) can be written in terms of the quantities

$$\begin{aligned}\phi_2 &= \beta_2(1 - b_2) \\ \psi_2 &= (1 - \beta_2)(1 - \lambda_2 b_2).\end{aligned}\tag{3.8}$$

In fact, we have

$$\begin{aligned}\kappa &= (1 - \phi_2 - \psi_2)^2 \\ \nu_1 &= 2\psi_2(1 - \phi_2 - \psi_2) \\ \nu_2 &= 2\phi_2(1 - \phi_2 - \psi_2) \\ \mu_{11} &= \psi_2^2 \\ \mu_{12} &= 2\phi_2\psi_2 \\ \mu_{22} &= \phi_2^2\end{aligned}\tag{3.9}$$

Moreover, it can easily be shown by analysis of (3.8) that any given non-negative pair (ϕ_2, ψ_2) such that $\phi_2 + \psi_2 \leq 1$ is compatible with any nonnegative value of λ_2 by suitable choice of β_2 and b_2 . Thus, under the submodel $H_{BB\&BC}$, λ_2 is non-identifiable in the same way as λ_1 , namely: λ_2 can take any value from 0 to ∞ while the values of the $\{\kappa\nu\mu\}$ probabilities are kept fixed.

On the other hand, some sets of $\{\kappa\nu\mu\}$ probabilities are clearly not compatible with all values of λ_2 . An interesting example of this is related to the submodels specified by the hypothesis $H: \lambda_2 = 1$. It is easy to see that this hypothesis implies that⁹

$$\frac{v_1}{v_1 + v_2} = \frac{\mu_{11} + \frac{1}{2}\mu_{12}}{\mu_{11} + \mu_{12} + \mu_{22}} \quad (3.10)$$

Consequently, any set of $\{\kappa\nu\mu\}$ parameters disagreeing with (3.10) is incompatible with the value $\lambda_2 = 1$. (Note that $H_{BB\&BC}$ implies, but is not implied by, the relationship (3.10)).

It is not clear at present how to precisely characterize those $\{\kappa\nu\mu\}$ configurations which determine λ_2 uniquely. It can be shown that if H_{BB} is true and H_{BC} is not, then λ_2 is determined uniquely by $\{\kappa\nu\mu\}$. Also, λ_2 is determined uniquely by $\{\kappa\nu\mu\}$ if H_{BC} is true and H_{BB} is not. It is not known what the situation is when H_{BB} and H_{BC} are both false.

⁹With the MBC λ model broadened as indicated in Footnote 8, the hypothesis $H: \lambda_2 = 1$ is equivalent to the hypothesis H_0^* of McCabe and Samuels [1], while Equation (3.10) is the hypothesis H_0 for which McCabe and Samuels derive the likelihood ratio test and the minimum modified χ^2 test.

It is not uncommon in parametric modeling to find that a parameter which is non-identifiable under a certain model may be identifiable under a more restrictive submodel. We have seen that the situation with the present model is just the reverse: it is under a submodel that the parameter λ_2 is not identifiable!

The practical consequence of this strange "partial non-identifiability" of λ_2 is that while some data sets may provide a great deal of information for estimating λ_2 , other data sets (for example, those fitting $H_{BB\&BC}$) provide no information whatsoever about λ_2 . It should be noted that this difficulty is not peculiar to the $MBC\lambda$ model--any model which contains $H_{BB\&BC}$ as a submodel (and any reasonable model would) will face the same kind of identifiability paradox.

3.3 Maximum Likelihood Estimation of all Parameters.

Because of the uncertainty in estimability of λ_2 , we start by considering λ_2 to be an arbitrary (preassigned) parameter, and derive estimates of the remaining parameters p_{12} , p_{22} , \bar{b}_2 , and \bar{b}_2^2 as functions of λ_2 . The likelihood function for the observations $(k, n_1, n_2, m_{11}, m_{12}, m_{22})$ is

$$L = \frac{(k + n_1 + n_2 + m_{11} + m_{12} + m_{22})!}{k! n_1! n_2! m_{11}! m_{12}! m_{22}!} \kappa^k v_1^{n_1} v_2^{n_2} u_{11}^{m_{11}} u_{12}^{m_{12}} u_{22}^{m_{22}} \quad (3.11)$$

The maximum likelihood estimates of the parameters under the $MBC\lambda$ model are determined by substituting (3.6) into (3.11) and maximizing L with respect to the parameters p_{12} , p_{22} (recalling that $p_{11} = 1 - 2p_{12} - p_{22}$), \bar{b}_2 and \bar{b}_2^2 . Note that, with λ_2 fixed, L factors into a part depending only on the branching parameters p_{ij} , and a part depending only on the cutting parameters \bar{b}_2 and \bar{b}_2^2 . Consequently, when λ_2 is preassigned, the branching

parameters and the cutting parameters are estimated independently of each other.

The maximization of L for fixed λ_2 may be carried out directly, using a suitable computer program¹⁰, or it may be done by setting the derivatives of $\log L$ with respect to the parameters equal to zero. The computations reported in Section 5 were done in the latter manner, using a computer subroutine to solve the resulting pairs of simultaneous nonlinear equations. These equations are given in the Appendix.

The amount of information about λ_2 present in a given data set can be judged by evaluating, for various preassigned values of λ_2 , the maximum of the likelihood function (3.11) with respect to p_{12} , p_{22} , $\overline{b_2}$ and $\overline{b_2^2}$. It is convenient to consider this maximum likelihood as a fraction of the unrestricted maximum L^* . L^* is the maximum of the likelihood function with respect to variation of the $\{\kappa\nu\mu\}$ parameters, subject only to the constraint $\kappa + \sum v_i + \sum \mu_{ij} = 1$; it is easy to show that this maximum L^* is achieved by the values

$$\kappa^* = k/x_1, \quad v_i^* = n_i/x_1, \quad \mu_{ij}^* = m_{ij}/x_1.$$

(Recall from Section 2 that $x_1 = k + n_1 + n_2 + m_{11} + m_{12} + m_{22}$.) The relative likelihood for a preassigned λ_2 value is, then¹¹,

$$R(\lambda_2) = \frac{\max L(p_{12}, p_{22}, \overline{b_2}, \overline{b_2^2})}{L^*}$$

¹⁰With this approach it is most efficient to maximize L/L^* (L^* is defined below) in order to avoid working with very small numbers.

¹¹ $R(\lambda_2)$ should not be confused with the vector R introduced earlier in this section.

$R(\lambda_2)$ is most conveniently computed as

$$R(\lambda_2) = \left(\frac{\kappa}{\kappa^*}\right)^k \left(\frac{v_1}{v_1^*}\right)^{n_1} \left(\frac{v_2}{v_2^*}\right)^{n_2} \left(\frac{\mu_{11}}{\mu_{11}^*}\right)^{m_{11}} \left(\frac{\mu_{12}}{\mu_{12}^*}\right)^{m_{12}} \left(\frac{\mu_{22}}{\mu_{22}^*}\right)^{m_{22}} \quad (3.13)$$

where κ , v_i and μ_{ij} are evaluated by substituting the maximum likelihood estimates of $\overline{b_2}$, b_2^2 and the p_{ij} into (3.6), and κ^* , v_i^* and μ_{ij}^* are as given above.

The maximum likelihood estimate of λ_2 is the value of λ_2 for which $R(\lambda_2)$ is maximum. The precision of this estimate is reflected in the rate at which $R(\lambda_2)$ declines away from the maximum. The asymptotic theory of maximum likelihood estimation (see, e.g., Wilks [4]) provides a guide to interpreting this precision, through the goodness-of-fit statistic¹²

$$\chi^2(\lambda_2) = -2 \log R(\lambda_2). \quad (3.14)$$

According to this theory, the interval of λ_2 values for which

$$\chi^2(\lambda_2) \leq 3.84$$

is an approximate (for large samples) 95% confidence interval for λ_2 .¹³

Approximate confidence intervals for the parameters p_{12} , p_{22} , $\overline{b_2}$ and b_2^2 can also be derived. However, these intervals will be broad if the interval for λ_2 is broad. Since this was the case for the data treated in Section 5, these computations have not been carried out.

¹²All logarithms used in this report are natural logarithms.

¹³The reason for this is that, if λ_2° is the true value of λ_2 , then the statistic $\chi^2(\lambda_2^\circ)$ has asymptotically (as the sample size goes to infinity) a chi-square distribution with 1 degree of freedom (d.f.) The value 3.84 is the 95th percentile of this chi-square distribution. If a confidence level other than 95% is desired, the percentile value must be adjusted accordingly.

3.4. Estimation and Testing for Fixed λ_2 .

When a given data set provides confidence intervals for λ_2 which are very broad or infinite, the only reasonable course is to treat λ_2 as an arbitrary parameter. The dependence of the other parameter estimates on the preassigned value of λ_2 can then be studied; the more cut branches there are in the data, the greater will be this dependence.

For fixed λ_2 , then, the maximum likelihood estimates of p_{12} , p_{22} , \bar{b}_2 and \bar{b}_2^2 can be computed as described in the preceding subsection. Approximate confidence intervals for these parameters could be obtained, but because of their dependence on the arbitrary value of λ_2 , this has not been done.

The maximum likelihood estimates \hat{p}_{12} and \hat{p}_{22} can be used to estimate the quantity N_3 , defined in [2] as the expected number of third-order branches per tree. In terms of the parametrization of Sections 2 and 3, $N_3 = N_2 \cdot 2(p_{12} + p_{22}) = 4\beta_1(p_{12} + p_{22})$, so that the maximum likelihood estimate of N_3 for given values of λ_1 and λ_2 is

$$\hat{N}_3 = 4\hat{\beta}_1(\hat{p}_{12} + \hat{p}_{22}) \quad (3.15)$$

where $\hat{\beta}_1$ is determined from Equation (2.4) or (2.4').

An approximate test for the goodness of fit of the data to the MBC λ model for a given λ_2 can be based on the statistic $\chi^2(\lambda_2)$ defined above. The fit is rejected as inadequate at the 5% significance level if $\chi^2(\lambda_2) > 3.84$ (See footnote 13). This test is valid asymptotically for large samples, so that its applicability to a given data set may be questionable. In this connection, it is worth noting that if, for a given data set, there is no value of λ_2 for which the fit is perfect (so that $\chi^2(\lambda_2) = 0$), doubt is cast upon the applicability of the goodness-of-fit test. This is because, with λ_2 varying, the MBC λ model has 5

independent parameters, and hence the same dimensionality as the unrestricted model under which the maximum L^* is achieved. Specifically, the parameter point $(\kappa, v_1, v_2, \mu_{11}, \mu_{12}, \mu_{22})$ is confined to a 5-dimensional hyperplane under the unrestricted model and to a 5-dimensional curvilinear subset of this hyperplane under the $MBC\lambda$ model. For small samples, the point $(\kappa^*, v_1^*, v_2^*, \mu_{11}^*, \mu_{12}^*, \mu_{22}^*)$ may not lie in this curvilinear subset, so that

$$\max L(\lambda_2, p_{12}, p_{22}, \bar{b}_2, \bar{b}_2^2) < L^*;$$

for samples large enough that the asymptotic theory is applicable, this situation cannot occur (unless the $MBC\lambda$ model is in fact invalid for all values of λ_2).

In addition to estimating the parameters of the $MBC\lambda$ model, it might be desired to estimate the branching probability β_2 under H_{BB} and the cutting probability b_2 under H_{BC} . If λ_2 is treated as an arbitrary parameter, then these two hypotheses are independent, and consequently the maximum likelihood estimates of β_2 under H_{BB} and b_2 under H_{BC} are the same as the maximum likelihood estimates of these parameters under $H_{BB\&BC}$. These latter are easily derived from Equations (3.8), (3.9), and (3.11). The resulting simple equations for estimating β_2 and b_2 will not be stated here, as they are given in more generality in Section 4.

The relative likelihood of H_{BB} is

$$R_{BB}(\lambda_2) = \frac{\max L(\beta_2, \bar{b}_2, \bar{b}_2^2)}{L^*}.$$

$R_{BB}(\lambda_2)$ may be computed from an expression like (3.13), with the parameters

κ , v_i and μ_{ij} evaluated by substituting the maximum likelihood estimates of β_2 , \bar{b}_2 and b_2^2 into (3.1) and then into (3.6). The goodness of fit of the data to H_{BB} for a fixed λ_2 can be measured by

$$\chi_{BB}^2(\lambda_2) = -2 \log R_{BB}(\lambda_2). \quad (3.16)$$

In a similar way, we may define for H_{BC}

$$R_{BC}(\lambda_2) = \frac{\max L(p_{12}, p_{22}, b_2)}{L^*}$$

and

$$\chi_{BC}^2(\lambda_2) = -2 \log R_{BC}(\lambda_2). \quad (3.17)$$

Again, R_{BC} may be computed from an expression like (3.13), with the parameters κ , v_i and μ_{ij} evaluated by substituting the maximum likelihood estimates of p_{12} , p_{22} and b_2 into (3.7) and then into (3.6).

Since H_{BB} and H_{BC} are independent for fixed λ_2 , the relative likelihood of the joint hypothesis $H_{BB \& BC}$ is

$$\begin{aligned} R_{BB \& BC} &= \frac{\max L(\beta_2, b_2)}{L^*} \\ &= \frac{R_{BB}(\lambda_2) \cdot R_{BC}(\lambda_2)}{R(\lambda_2)} \end{aligned}$$

and, correspondingly,

$$\chi_{BB \& BC}^2 = \chi_{BB}^2(\lambda_2) + \chi_{BC}^2(\lambda_2) - \chi^2(\lambda_2)$$

Note that, because of the non-identifiability of λ_2 under $H_{BB \& BC}$, the goodness-

of-fit statistic $\chi_{BB\&BC}^2$ does not depend on λ_2 .

Approximate tests of H_{BB} , H_{BC} and $H_{BB\&BC}$ against the unrestricted model can be based on these χ^2 statistics. $\chi_{BB}^2(\lambda_2)$ has (asymptotically for large samples) a chi-square distribution with 2 d.f. if H_{BB} is true for the specified value of λ_2 . Similarly for $\chi_{BC}^2(\lambda_2)$. If $H_{BB\&BC}$ is true, $\chi_{BB\&BC}^2$ has (asymptotically for large samples) a chi-square distribution with 3 d.f.

For a fixed value of λ_2 , H_{BB} , H_{BC} , and $H_{BB\&BC}$ can be approximately tested against the $MBC\lambda$ model with the statistics

$$\chi_{BB-MBC}^2(\lambda_2) = \chi_{BB}^2(\lambda_2) - \chi^2(\lambda_2), \quad (3.18)$$

$$\chi_{BC-MBC}^2(\lambda_2) = \chi_{BC}^2(\lambda_2) - \chi^2(\lambda_2), \quad (3.19)$$

and
$$\chi_{BB\&BC-MBC}^2(\lambda_2) = \chi_{BB\&BC}^2 - \chi^2(\lambda_2), \quad (3.20)$$

respectively. These statistics have (asymptotically for large samples) χ^2 distributions with 1, 1, and 2 d.f., respectively, if the corresponding null hypotheses are true. In interpreting these tests, it is necessary to keep in mind that the value of λ_2 is forced to remain the same under the null and alternative hypotheses. The results may not be exactly the same as if λ_2 were allowed to vary during the maximization process.

All of the above χ^2 tests, both against the unrestricted model and against the $MBC\lambda$ model, are only approximately valid for large samples; the cautions mentioned above in connection with testing the fit of the $MBC\lambda$ model should be observed.

4. EXTENSION OF THE MODELS TO ALL ORDERS OF BRANCHING

In order to estimate the branching parameters for branches of order 3 and higher, we consider the extension of the MBC λ model of Section 3 to an arbitrary order k of branching. Attempts to extend the full MBC λ model have led to intractable equations. However, the equations simplify very considerably if a binomial branching assumption, analogous to H_{BB} , is made. In this section we derive the equations for estimating the k^{th} -order branching probabilities under an extended MBC λ model with a binomial branching assumption incorporated. Estimation of the cutting probabilities will be considered only under the additional hypothesis of binomial cutting.

The models will first be formulated for each of the T trees separately. For the i^{th} -tree, define the following random variables:

$$\begin{aligned} X_k^{(i)} &= \text{number of bifurcating uncut } k^{\text{th}}\text{-order branches} \\ Y_k^{(i)} &= \text{number of terminal uncut } k^{\text{th}}\text{-order branches} \\ Z_k^{(i)} &= \text{number of cut } k^{\text{th}}\text{-order branches.} \end{aligned}$$

All the branches mentioned in this definition are assumed not to have been excised by cuts on branches of order $(k-1)$ or lower. Thus

$$V_k^{(i)} = X_k^{(i)} + Y_k^{(i)} + Z_k^{(i)}$$

is the total number of observable k^{th} -order branches on the i^{th} tree.

We treat the various orders of branching separately by conditioning, at each level, on $V_k^{(i)}$. For each fixed k and i , then, consider the following assumptions, conditional on the event $V_k^{(i)} = v_k^{(i)}$:

1. Binomial Branching Assumption. The $v_k^{(i)}$ branches bifurcate independently of each other, each with bifurcation probability β_k .

2. Modified Binomial Cutting Assumption. There is a vector random variable $R_k^{(i)}$, such that conditional on $R_k^{(i)} = \underline{r}$, the $v_k^{(i)}$ branches are cut independently of each other, with cutting probabilities $t_k(\underline{r})$ for terminal branches and $b_k(\underline{r})$ for bifurcating branches. The random variable R_k has distribution function¹⁴ $F_k(\underline{r}) = P\{R_k \leq \underline{r}\}$.

3. Constant-Ratio Assumption. There is a constant, λ_k , such that, for all \underline{r} ,

$$\frac{t_k(\underline{r})}{b_k(\underline{r})} = \lambda_k.$$

4. Binomial Cutting Assumption. The $v_k^{(i)}$ branches are cut or not cut independently of each other, with cutting probabilities t_k for terminal branches and b_k for bifurcating branches. This is a special case of assumption (2) with $t_k(\underline{r}) \equiv t_k$ and $b_k(\underline{r}) \equiv b_k$.

For brevity, we will call the model defined by assumptions (1), (2), and (3) the binomial branching (BB) model. The submodel defined by adding assumption (4) will be called the binomial branching and cutting (BB&BC) model.

For reasons perfectly analogous to those discussed in Section 3, the parameter λ_k is not identifiable under the BB&BC model, and consequently it may or may not be estimable under the BB model. Throughout this section, we shall treat λ_k as an arbitrary pre-specified parameter.

Under assumptions (1) and (2), the probability distribution $P_k^{(i)}$ for the configuration of the k^{th} -order branches of the i^{th} tree is

¹⁴The inequality $R_k \leq \underline{r}$ means coordinate-wise inequality.

$$P_k^{(i)} = P\{X_k^{(i)} = x_k^{(i)}, Y_k^{(i)} = y_k^{(i)}, Z_k^{(i)} = z_k^{(i)} | V_k^{(i)} = v_k^{(i)}\} =$$

$$\frac{v_k^{(i)}!}{x_k^{(i)}! y_k^{(i)}! z_k^{(i)}!} \int \{\beta_k [1-b_k(r)]\}^{x_k^{(i)}} \{[1-\beta_k][1-t_k(r)]\}^{y_k^{(i)}} \{\beta_k b_k(r) + (1-\beta_k)t_k(r)\}^{z_k^{(i)}} dF_k(r)$$
(4.1)

Upon incorporating assumption (3), Equation (4.1) reduces to the form

$$P_k^{(i)} = C_k^{(i)} \beta_k^{x_k^{(i)}} (1-\beta_k)^{y_k^{(i)}} [\beta_k + (1-\beta_k)\lambda_k]^{z_k^{(i)}}, \quad (4.2)$$

where $C_k^{(i)}$ (which includes the integral with respect to F_k) does not depend on β_k .

We now assume that the T trees comprising the data are independent realizations of the model, so that the joint conditional distribution P_k of the configuration of all the k^{th} -order branches are all the trees is, under the BB model,

$$P_k = P\left\{ \bigcap_i (X_k^{(i)} = x_k^{(i)}, Y_k^{(i)} = y_k^{(i)}, Z_k^{(i)} = z_k^{(i)}) \mid \bigcap_i V_k^{(i)} = v_k^{(i)} \right\}$$

$$= \prod_i P_k^{(i)}$$

$$= C_k \beta_k^{x_k} (1-\beta_k)^{y_k} [\beta_k + (1-\beta_k)\lambda_k]^{z_k} \quad (4.3)$$

where $x_k = \sum_i x_k^{(i)}$, $y_k = \sum_i y_k^{(i)}$, $z_k = \sum_i z_k^{(i)}$, and $C_k = \prod_i C_k^{(i)}$ does not depend on β_k .

Equation (4.3) reveals two facts about the estimation of β_k under the BB model. First, the statistics x_k , y_k , and z_k are sufficient for the estimation of β_k - that is, the distribution of the bifurcating, terminal and cut branches among the T trees need not be specified. Second, since the dependence of P_k upon β_k is of the same form with or without the incorporation of the

Binomial Cutting Assumption (4), the maximum likelihood estimator $\hat{\beta}_k$ is the same under the BB model as under the BB&BC model.

We proceed, then, to estimate β_k and b_k under the BB&BC model. Incorporating assumption (4) into Equations (4.1), (4.2) and (4.3), P_k can be written

$$P_k = C'_k \phi_k^{x_k} \psi_k^{y_k} (1 - \phi_k - \psi_k)^{z_k} \quad (4.4)$$

where

$$\begin{aligned} \phi_k &= \beta_k(1-b_k) \\ \psi_k &= (1-\beta_k)(1-\lambda_k b_k) \end{aligned} \quad (4.5)$$

and C'_k does not depend on β_k or on b_k . The maximum likelihood estimates of ϕ_k and ψ_k are easily seen to be

$$\hat{\phi}_k = x_k/v_k \quad (4.6)$$

and

$$\hat{\psi}_k = y_k/v_k,$$

where $v_k = \sum_i v_k^{(i)} = x_k + y_k + z_k$. Combining (4.5) and (4.6) gives the following equations for the maximum likelihood¹⁵ estimates $\hat{\beta}_k$, \hat{b}_k , and \hat{t}_k :

$$\hat{\beta}_k^2(1 - \lambda_k)v_k - \hat{\beta}_k[x_k + z_k - \lambda_k(x_k + v_k)] - \lambda_k x_k = 0 \quad (4.7)$$

$$\hat{b}_k = 1 - x_k/\hat{\beta}_k v_k \quad (4.8)$$

$$\hat{t}_k = \lambda_k \hat{b}_k \quad (4.9)$$

¹⁵The fact that P_k is conditioned on V_k does not affect the maximum likelihood estimators; see Footnote 3.

It is easy to show that if both x_k and y_k are nonzero, then for each λ_k , $0 \leq \lambda_k < \infty$, Equation (4.7) has a unique¹⁶ solution $\hat{\beta}_k$ between 0 and 1.

In the case $\lambda_k = \infty$, it can be seen by obvious modifications of the derivation that (4.7), (4.8), and (4.9) must be replaced by:

$$\hat{\beta}_k = x_k / v_k \quad (4.7')$$

$$\hat{b}_k = 0 \quad (4.8')$$

$$\hat{t}_k = \frac{z_k}{y_k + z_k} \quad (4.9')$$

We note in passing that the correction for cutting proposed by Smit, Uylings, and Veldmaat-Wansink [3] is equivalent to applying Equations (4.7), (4.8), and (4.9) with $\lambda_k = 2$.

Equations (4.7) through (4.9') are valid under the BB&BC model. Equations (4.7) and (4.7') are valid under the BB model as well. For $k = 1$, the BB&BC model is trivially valid; in this case, Equations (4.7) through (4.9') are equivalent to (2.4) through (2.5'). For $k = 2$, as was mentioned in Section 3, the estimators (4.8), (4.9), (4.8') and (4.9') of the cutting parameters are valid under H_{BC} , regardless of whether H_{BB} is true. Of course, if H_{BB} is false, the quantity $\hat{\beta}_2$ of (4.7) cannot be interpreted as a bifurcation probability; nevertheless, substitution of the solution $\hat{\beta}_2$ into (4.8) gives the maximum likelihood estimate of b_2 .

The reader will recall that the notation for observed frequencies used in Section 3 is more complicated than that of the present section. This is because the statistics x_2 , y_2 , and z_2 are not sufficient for estimation of cutting and

¹⁶If either x_k or y_k is zero, the maximum likelihood estimates are not unique for certain values of λ_k .

branching parameters under the full MBC λ model of Section 3. The relationship between the two notations is

$$\begin{aligned}
 x_2 &= n_2 + m_{12} + 2m_{22} \\
 y_2 &= n_1 + 2m_{11} + m_{12} \\
 z_2 &= 2k + n_1 + n_2 \\
 v_2 &= 2(k + n_1 + n_2 + m_{11} + m_{12} + m_{22})
 \end{aligned}
 \tag{4.10}$$

In [2], we consider the quantity N_k , defined as the mean number of k^{th} -order branches per tree. Of course, $N_1 = 1$. Under the BB or BB&BC model,

$$N_k = 2^{k-1} \beta_1 \beta_2 \dots \beta_{k-1},$$

so that the maximum likelihood estimates of the N_k under these models are obtained from the $\hat{\beta}_k$ as

$$\begin{aligned}
 \hat{N}_2 &= 2\hat{\beta}_1 \\
 \hat{N}_3 &= 4\hat{\beta}_1 \hat{\beta}_2 \\
 &\vdots \\
 \hat{N}_k &= 2^{k-1} \hat{\beta}_1 \hat{\beta}_2 \dots \hat{\beta}_{k-1}
 \end{aligned}$$

5. ANALYSIS OF DATA

5.1 Numerical Results.

In this section we present the data used in [2], and some of the results of applying the methods of the previous sections to this data. The main results are presented in [2].

The data was collected by Professor Paul D. Coleman, as described in [2]. Observations were made on four sets of nerve cells. Because the branching parameters appeared to depend upon the number of dendritic trees borne on the same nerve cell, each of the sets of cells was divided into three groups: cells bearing 4 or fewer trees, cells bearing 5 trees, and cells bearing 6 or more trees. Thus it is to be hoped that each group of trees represents approximately a sample from a homogeneous population, as required by the models of the previous sections.

Table I gives an abbreviated name for and brief description of each of the 12 groups, together with the numbers of dendritic trees and of nerve cells in each group. Table II gives, for each group, the observed frequencies of the tree types defined by first- and second-order branching patterns; the notation is that of Sections 2 and 3 of this report. Table III gives the observed frequencies of bifurcating, terminal, and cut branches of orders 3 and 4; the notation is that of Section 4 of this report. The frequencies x_5 were zero for all groups; the frequencies y_5 and z_5 were small and are not reported here.

Table IV shows values of the statistic $\chi^2(\lambda_2)$, which measures (see Equation (3.14)) the goodness of fit of the $MBC\lambda$ model, for selected values of λ_2 . It can be seen that the shape of the function $\chi^2(\lambda_2)$, as λ_2 varies from 0 to ∞ , is not the same for all the groups: for some groups it is

TABLE I. DESCRIPTION OF DATA GROUPS

Abbreviated Name of Group	Type of Cell	Number of Trees per Cell	Number of Cells	Number of Trees
SC \leq 4	layer IV stellate from control cats	\leq 4	14	54
SC 5		5	17	85
SC \geq 6		\geq 6	14	93
				<u>232</u>
SD \leq 4	layer IV stellate from dark- reared cats	\leq 4	12	44
SD 5		5	20	100
SD \geq 6		\geq 6	13	82
				<u>226</u>
P1 \leq 4	layer V pyramidal from strain S1 rats	\leq 4	21	78
P1 5		5	16	80
P1 \geq 6		\geq 6	7	44
P3 \leq 4	layer V pyramidal from strain S3 rats	\leq 4	21	77
P3 5		5	15	75
P3 \geq 6		\geq 6	8	50

TABLE II. FIRST- AND SECOND-ORDER DATA

Group	y_1	z_1	k	n_1	n_2	m_{11}	m_{12}	m_{22}
SC \leq 4	0	0	5	3	9	9	16	12
SC 5	3	1	3	11	15	16	24	12
SC $>$ 6	11	5	12	10	14	22	13	6
SD \leq 4	2	2	4	7	5	5	9	10
SD 5	2	7	8	23	12	16	22	10
SD $>$ 6	8	7	13	13	12	15	10	4
P1 \leq 4	8	0	2	7	8	25	19	11
P1 5	5	4	6	15	7	14	24	5
P1 $>$ 6	6	3	2	5	5	9	11	3
P3 \leq 4	0	3	3	4	7	20	26	14
P3 5	6	3	1	13	8	22	15	7
P3 $>$ 6	3	2	3	5	5	16	14	2

TABLE II. FIRST- AND SECOND-ORDER DATA

Group	y_1	z_1	k	n_1	n_2	m_{11}	m_{12}	m_{22}
SC ≤ 4	0	0	5	3	9	9	16	12
SC 5	3	1	3	11	15	16	24	12
SC > 6	11	5	12	10	14	22	15	6
SD ≤ 4	2	2	4	7	5	5	9	10
SD 5	2	7	8	23	12	16	22	10
SD > 6	8	7	13	15	12	15	10	4
P1 ≤ 4	8	0	2	7	8	23	19	11
P1 5	5	4	6	15	7	14	24	5
P1 ≥ 6	6	3	2	5	5	9	11	3
P3 ≤ 4	0	3	3	4	7	20	26	14
P3 5	6	3	1	13	8	22	15	7
P3 > 6	3	2	3	5	5	16	14	2

TABLE III. THIRD- AND FOURTH-ORDER DATA

Group	x_3	y_3	z_3	x_4	y_4	z_4
SC \leq 4	18	55	25	4	19	13
SC 5	24	75	27	7	33	8
SC \geq 6	14	41	23	1	18	9
SD \leq 4	15	40	13	3	19	8
SD 5	18	59	31	4	23	9
SD \geq 6	9	28	23	0	9	9
P1 \leq 4	19	58	21	2	26	10
P1 5	13	53	16	1	17	8
P1 \geq 6	7	29	8	1	11	2
P3 \leq 4	16	83	23	4	21	7
P3 5	13	51	10	4	19	3
P3 \geq 6	9	30	7	1	14	3

TABLE IV. $\chi^2(\lambda_2)$

Group	λ_2										
	0	.01	.1	.5	1	2	3	4	10	100	∞
SC \leq 4	2.88	2.89	2.99	2.88	1.91	.49	.15	.06	.00	.02	.02
SC \leq 5	.13	.14	.22	.68	1.09	1.29	1.28	1.25	1.15	1.06	1.05
SC \geq 6	1.68	1.75	2.51	5.81	5.64	3.23	2.45	2.12	1.59	1.33	1.31
SD \leq 4	.39	.39	.47	.87	1.31	1.29	1.08	.94	.66	.50	.48
SD \leq 5	.19	.20	.28	.74	.94	.78	.66	.60	.48	.41	.40
SD \geq 6	.55	.57	.81	2.15	2.01	.98	.70	.58	.41	.32	.31
P1 \leq 4	.02	.03	.10	.66	1.11	1.41	1.49	1.52	1.56	1.57	1.57
P1 \leq 5	1.39	1.35	.99	.01	.47	1.42	1.80	1.99	2.31	2.48	2.49
P1 \geq 6	.71	.72	.78	.81	.56	.29	.19	.15	.08	.05	.05
P3 \leq 4	1.99	2.00	2.09	2.00	1.40	.65	.40	.29	.14	.08	.08
P3 \leq 5	1.58	1.53	1.10	.36	.19	.79	1.10	1.27	1.59	1.78	1.80
P3 \geq 6	3.33	3.37	3.59	3.16	1.83	.75	.45	.33	.15	.08	.08

monotone; for some groups it appears to have a single local minimum for some intermediate value of λ_2 ; for some groups it appears to have a single local maximum. Note also that, for several of the groups, the minimum value of $\chi^2(\lambda_2)$ is not zero, so that there is no value of λ_2 for which the MBC λ model fits perfectly. On the other hand, the model fits moderately well for almost all values of λ_2 ; the fit is rejected by the asymptotic test at the 5% significance level ($\chi^2 > 3.84$) only for one data group ($SC \geq 6$), and only for a rather small range of λ_2 . As was mentioned in Section 3, the fact that the minimum χ^2 value is not zero suggests that the asymptotic theory may not be very accurate for this data; nevertheless, it appears safe to conclude from the flatness of the $\chi^2(\lambda_2)$ functions (and also from the fact that they have various shapes) that this data contains little or no information about the value of λ_2 .

Table V shows selected values of the statistic $\chi_{\text{BB-MBC}}^2(\lambda_2)$ (defined in Equation (3.18)), which measures the goodness-of-fit of the data to the binomial branching hypothesis H_{BB} . The direction of deviation from H_{BB} is also indicated--cases where the dependence between bifurcation of sister branches is negative, so that $\hat{p}_{22} < (\hat{p}_{12} + \hat{p}_{22})^2$, are marked by an asterisk (*) in Table V; in the remaining cases the inequality is reversed, showing positive dependence. Note that the only cases of substantial deviations (recall that the asymptotic critical value is 2.84)¹⁷, are associated with positive dependence. Furthermore, these cases occur largely in the somewhat heterogeneous groups--i.e., those having ≤ 4 or ≥ 6 trees per cell. It is well known that the effect of heterogeneity within a group is a tendency toward

¹⁷Of course, H_{BB} should not be tested within the MBC λ model for the data group $SC \geq 6$ in the range of λ_2 values for which the MBC λ model was rejected by $\chi^2(\lambda_2)$. However, this has no effect on the general conclusions drawn about H_{BB} . The same remark applies to the test of H_{BC} considered below.

TABLE V. $\chi^2_{BB-MBC}(\lambda_2)$

Group	λ_2										
	0	.01	1.	.5	1	2	3	4	10	100	∞
SC \leq 4	2.54	2.51	2.22	1.24	.65	.33	.26	.24	.23	.23	.24
SC \leq 5	.96	.96	.90	.59	.28	.06	.02	.00	.00	.01	.01
SC \geq 6	12.10	12.01	11.15	6.42	2.88	1.16	.80	.67	.48	.41	.40
SD \leq 4	.31	.31	.36	.66	1.18	1.98	2.39	2.62	3.01	3.22	3.25
SD \leq 5	.02*	.01*	.01*	.03	.24	.57	.71	.78	.90	.97	.98
SD \geq 6	5.33	5.31	5.02	3.01	1.20	.45	.33	.28	.22	.20	.20
P1 \leq 4	4.66	4.65	4.56	3.89	3.19	2.51	2.22	2.07	1.79	1.62	1.60
P1 \leq 5	.74*	.76*	.94*	1.41*	1.21*	.79*	.61*	.52*	.38*	.30*	.29*
P1 \geq 6	.25	.25	.18	.01	.01*	.07*	.09*	.10*	.12*	.13*	.13*
P3 \leq 4	2.60	2.56	2.26	1.39	.95	.68	.59	.56	.50	.47	.47
P3 \leq 5	1.38	1.41	1.69	2.35	2.25	1.86	1.65	1.54	1.33	1.20	1.18
P3 \geq 6	.89	.85	.52	.00*	.20*	.41*	.48*	.51*	.55*	.57*	.57*

apparent positive dependence. On balance, it appears that the data are quite consistent with the hypothesis H_{BB} , with any apparent tendency toward positive dependence being an artifact caused by the heterogeneity.

In any case, a slight dependence (real or spurious) between sister branches with respect to bifurcation would not be expected to have much effect on the estimation of N_3 , since this is a mean. From Equation (3.15) we see that the effect would be felt through the estimation of the quantity $(p_{12} + p_{22})$. Under H_{BB} , this quantity would be replaced by β_2 . It can be seen from Equation (4.7) and Equation (A1) and (A2) of the Appendix that $\hat{p}_{12} + \hat{p}_{22} = \hat{\beta}_2$ if $\lambda_2 = 0$; analogously, these estimates are equal if $\lambda_2 = \infty$. Table VI shows the comparison between the estimates for several intermediate values of λ_2 . The differences are seen to be relatively quite small. It appears likely, therefore, that use of the BB assumption for estimation of N_4, N_5 , etc. (as in Section 4) is not unreasonable.

Table VII provides an informative companion to Table V. In Table VII are presented selected values of $\chi_{BC-MBC}^2(\lambda_2)$ (defined in Equation (3.19)), which may be used to test the binomial cutting hypothesis H_{BC} within the $MBC\lambda$ model. Because of the nature of the experiment, H_{BC} is very likely a priori to be false in the direction of positive dependence: i.e., $b_2^2 > (\bar{b}_2)^2$. The entries in Table VII for which the corresponding estimates do not satisfy this inequality are marked with an asterisk (*). Note that there are very few such entries, and that they are associated with small χ^2 values. In testing H_{BC} , it is appropriate to use a one-sided alternative of positive dependence because of the a priori consideration mentioned above; thus, the asymptotic test will reject H_{BC} at the 5% significance level if the entry is not marked by an asterisk and if the χ^2 value exceeds 2.71.

TABLE VI ($\hat{p}_{12} + \hat{p}_{22}$) AND $\hat{\beta}_2$

λ_2	.5		1		2	
Group	$\hat{p}_{12} + \hat{p}_{22}$	$\hat{\beta}_2$	$\hat{p}_{12} + \hat{p}_{22}$	$\hat{\beta}_2$	$\hat{p}_{12} + \hat{p}_{22}$	$\hat{\beta}_2$
SC \leq 4	.612	.608	.573	.570	.528	.526
SC = 5	.526	.525	.486	.485	.446	.446
SC \geq 6	.463	.445	.381	.368	.315	.311
SD \leq 4	.613	.616	.560	.567	.504	.511
SD = 5	.478	.478	.411	.412	.357	.358
SD \geq 6	.476	.466	.369	.361	.290	.288
P1 $<$ 4	.435	.432	.409	.405	.385	.382
P1 = 5	.433	.434	.382	.380	.339	.337
P1 \geq 6	.436	.436	.392	.393	.357	.358
P3 \leq 4	.490	.487	.467	.466	.446	.445
P3 = 5	.375	.375	.341	.339	.315	.313
P3 \geq 6	.347	.347	.309	.311	.284	.285

TABLE VII χ^2 BC-MBC (λ_2)

λ_2 / Group	0	.01	.1	.5	1	2	3	4	10	100	∞
SC \leq 4	1.83	1.85	2.05	3.13	4.68	6.42	6.83	6.95	7.02	6.99	6.99
SC = 5	.29*	.29*	.26*	.12*	.01*	.03	.09	.14	.24	.31	.32
SC \geq 6	.32	.33	.44	1.86	5.58	9.71	10.84	11.32	12.02	12.36	12.40
SD \leq 4	3.29	3.28	3.16	2.45	1.50	.72	.51	.42	.31	.26	.25
SD = 5	1.18	1.17	1.10	.62	.19	.03	.01	.00	.00	.00	.00
SD \geq 6	.22	.22	.27	.94	2.90	4.66	5.08	5.24	5.47	5.58	5.59
PI \leq 4	.09	.09	.10	.23	.47	.85	1.06	1.18	1.42	1.58	1.60
PI = 5	1.04	1.06	1.25	1.75	1.49	.96	.76	.66	.49	.40	.39
PI \geq 6	.00	.00	.01	.14	.38	.61	.68	.72	.77	.79	.79
P3 \leq 4	1.72	1.74	1.95	2.92	3.96	4.98	5.31	5.46	5.66	5.75	5.76
P3 = 5	.32*	.34*	.50*	.90*	.85*	.64*	.54*	.48*	.37*	.31*	.30*
P3 \geq 6	.04	.05	.16	1.10	2.23	3.10	3.33	3.43	3.56	3.61	3.61

(the 90th percentile of the chi-square distribution with 1 d.f.) With this criterion, H_{BC} is rejected by 5 of the 12 data groups for almost all values $\lambda_2 \geq 1$ (there is a minor exception in P3 ≥ 6). It is reassuring to thus demonstrate that the statistic χ_{BC-MBC}^2 is sufficiently sensitive to detect, at least for some of the data groups, the departure from H_{BC} which we know must exist.

In [2] the dependence of \hat{N}_3 on the number of trees per cell is discussed. Table VIII shows values of \hat{N}_3 (calculated from Equation (3.15)) for selected values of λ_1 and λ_2 . Because \hat{N}_3 is a non-increasing function of λ_1 , the behavior of \hat{N}_3 for values of λ_1 between 0 and ∞ can easily be inferred from Table VIII. It can be seen that the value of λ_1 has relatively little effect upon \hat{N}_3 (this is because only a few first-order branches were cut). For the SC, SD, and P3 data sets, the monotone decreasing pattern of dependence of \hat{N}_3 on the number of trees per cell can be seen to hold for all λ_1 and λ_2 such that $0 \leq \lambda_1 \leq \infty$ and $.5 \leq \lambda_2 \leq \infty$.¹⁸ This conclusion holds even if the λ_1 and λ_2 values depend upon the number of trees per cell, as long as the dependence is not too strong. For the P1 data set, the monotone decreasing pattern is less pronounced, and disappears as λ_2 drops a little below 1.0.

In [2], we present the estimates \hat{N}_2 , \hat{N}_3 , \hat{N}_4 , and \hat{N}_5 , calculated with all $\lambda_k = 2$. We also discuss in [2] the quantities $\hat{W}_k = \hat{N}_k(T/C)$, where T/C is the mean number of trees per cell in a given data group. Thus \hat{W}_k is the estimated number of k^{th} -order branches per cell for the data group. Tables IX through XIV show the values of the \hat{N}_k and the \hat{W}_k for $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$,

¹⁸ Recall that there is strong presumptive evidence that $\lambda_2 \geq 1$.

TABLE VIII. \hat{N}_5 FOR SELECTED VALUES OF λ_1 AND λ_2

Group	$\lambda_2 = .5$		$\lambda_2 = 1$		$\lambda_2 = 2$		$\lambda_2 = 10$		$\lambda_2 = \infty$	
	$\lambda_1 = 0$	$\lambda_1 = \infty$	$\lambda_1 = 0$	$\lambda_1 = \infty$	$\lambda_1 = 0$	$\lambda_1 = \infty$	$\lambda_1 = 0$	$\lambda_1 = \infty$	$\lambda_1 = 0$	$\lambda_1 = \infty$
SC ≤ 4	2.45	2.44	2.29	2.29	2.11	2.11	1.88	1.88	1.82	1.81
SC ≥ 5	1.93	2.00	1.88	1.85	1.72	1.70	1.72	1.52	1.50	1.48
SC ≥ 6	1.63	1.54	1.34	1.26	1.11	1.04	.93	.88	.89	.84
SD ≤ 4	2.34	2.23	2.14	2.04	1.92	1.83	1.69	1.61	1.63	1.55
SD ≥ 5	1.87	1.74	1.61	1.50	1.40	1.30	1.21	1.12	1.16	1.08
SD ≥ 6	1.72	1.55	1.48	1.21	1.05	.95	.85	.77	.81	.73
PI ≤ 4	1.56	1.56	1.47	1.47	1.38	1.38	1.28	1.29	1.26	1.26
PI ≥ 5	1.62	1.54	1.43	1.36	1.27	1.20	1.12	1.06	1.08	1.03
PI ≥ 6	1.51	1.39	1.36	1.25	1.23	1.14	1.12	1.03	1.08	1.00
P3 ≤ 4	1.96	1.88	1.87	1.80	1.78	1.71	1.68	1.61	1.65	1.58
P3 ≥ 5	1.38	1.32	1.26	1.20	1.16	1.11	1.06	1.01	1.03	.99
P3 ≥ 6	1.30	1.25	1.16	1.11	1.07	1.02	.98	.94	.96	.92

.5, 1, 2, 4, and ∞ . In addition, these tables show a quantity denoted " W_{sum} ". This quantity is equal to $W_2 + W_3 + W_4 + W_5$ for the SC and SD data sets, but is equal to $W_3 + W_4 + W_5$ for the P1 and P3 data sets. The following phenomena are noted in [2]: as the number of trees per cell increases, (i) \hat{N}_3 , \hat{N}_4 , and \hat{N}_5 decrease monotonely within each data set (with a few minor exceptions); and (ii) W_{sum} is approximately constant within each data set except P1. The extent to which these phenomena persist as the assumed common value of the λ_k varies may be judged by the reader. As an aid to perusal of the tables, data sets for which the monotonicity of \hat{N}_k fails are marked by an asterisk, and the maximum percentage variation of W_{sum} within each data set is given in parentheses.

TABLE IX. \hat{N}_k and \hat{W}_k for $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$

Group	\hat{N}_2	\hat{N}_3	\hat{N}_4	\hat{N}_5	\hat{W}_2	\hat{W}_3	\hat{W}_4	\hat{W}_5	\hat{W}_{sum}
SC ≤ 4	2.00	2.63	2.51	2.18	7.7	10.1	8.9	8.4	35.2
SC ≥ 5	1.95	2.26	1.83	1.14	9.6	11.3	9.2	5.7	35.8
SC ≥ 6	1.76	1.99	1.89	1.35	11.7	13.2	12.6	9.0	46.5
			*	*					(32%)
SD ≤ 4	1.91	2.58	2.12	1.56	7.0	9.4	7.8	5.7	29.9
SD ≥ 5	1.96	2.26	2.05	1.44	9.8	11.3	10.3	7.2	38.6
SD ≥ 6	1.80	2.18	2.33	0	11.4	13.8	14.7	0	39.8
	*		*						(33%)
P1 ≤ 4	1.79	1.74	1.42	.90	6.7	6.5	5.3	3.3	15.1
P1 ≥ 5	1.88	1.98	1.40	.97	9.4	9.9	7.0	4.8	21.8
P1 ≥ 6	1.73	1.78	1.21	.52	10.9	11.2	7.6	3.3	22.0
	*	*		*					(46%)
P3 ≤ 4	2.00	2.11	1.35	.93	7.3	7.7	4.9	3.4	16.1
P3 ≥ 5	1.84	1.67	1.04	.56	9.2	8.4	5.2	2.8	16.4
P3 ≥ 6	1.88	1.63	1.13	.50	11.8	10.2	7.1	3.1	20.4
	*		*						(27%)

TABLE X. \hat{N}_k and \hat{W}_k for $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = .5$

Group	\hat{N}_2	\hat{N}_3	\hat{N}_4	\hat{N}_5	\hat{W}_2	\hat{W}_3	\hat{W}_4	\hat{W}_5	\hat{W}_{sum}
SC \leq 4	2.00	2.45	1.48	.77	7.7	9.4	5.7	3.0	25.8
SC = 5	1.93	2.03	1.16	.47	9.6	10.2	5.8	2.3	27.9
SC \geq 6	1.76	1.63	1.05	.18	11.7	10.8	7.0	1.2	30.7 (19%)
SD \leq 4	1.91	2.34	1.45	.53	7.0	8.6	5.3	1.9	22.8
SD = 5	1.96	1.87	1.12	.41	9.8	9.4	5.6	2.0	26.8
SD \geq 6	1.80 *	1.71	1.19 *	0	11.3	10.8	7.5	0	29.6 (30%)
P1 \leq 4	1.79	1.56	.91	.18	6.7	5.8	3.4	.7	9.8
P1 = 5	1.87	1.62	.75	.13	9.4	8.1	3.8	.7	12.5
P1 \geq 6	1.72 *	1.50 *	.68	.13	10.8	9.4	4.3	.8	14.5 (47%)
P3 \leq 4	2.00	1.96	.75	.30	7.3	7.2	2.8	1.1	11.0
P3 = 5	1.84	1.38	.62	.24	9.2	6.9	3.1	1.2	11.2
P3 \geq 6	1.88 *	1.30	.67 *	.11	11.7	8.1	4.2	.7	13.0 (18%)

TABLE XI. \hat{N}_k and \hat{W}_k for $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 1$.

Group	\hat{N}_2	\hat{N}_3	\hat{N}_4	\hat{N}_5	\hat{W}_2	\hat{W}_3	\hat{W}_4	\hat{W}_5	\hat{W}_{sum}
SC ≤ 4	2.00	2.29	1.13	.39	7.7	8.8	4.4	1.5	22.4
SC ≥ 5	1.93	1.87	.91	.32	9.6	9.4	4.5	1.6	25.2
SC ≥ 6	1.75	1.33	.68	.07	11.6	8.9	4.5	.5	25.5 (14%)
SD ≤ 4	1.90	2.13	1.16	.32	7.0	7.8	4.3	1.2	20.2
SD ≥ 5	1.96	1.61	.75	.22	9.8	8.0	3.8	1.1	22.7
SD ≥ 6	1.79 *	1.32	.64	0	11.3	8.3	4.0	0	23.6 (17%)
P1 ≤ 4	1.79	1.47	.72	.10	6.7	5.4	2.7	.4	8.5
P1 ≥ 5	1.87	1.43	.56	.06	9.3	7.1	2.8	.3	10.2
P1 ≥ 6	1.71 *	1.34	.52	.09 *	10.7	8.4	3.3	.5	12.2 (44%)
P3 ≤ 4	2.00	1.87	.60	.19	7.3	6.9	2.2	.7	9.8
P3 ≥ 5	1.83	1.25	.51	.18	9.2	6.3	2.5	.9	9.7
P3 ≥ 6	1.88 *	1.16	.53 *	.07	11.7	7.2	3.3	.4	11.0 (14%)

TABLE XII. \hat{N}_k and \hat{W}_k for $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 2$.

Group	\hat{N}_2	\hat{N}_3	\hat{N}_4	\hat{N}_5	\hat{W}_2	\hat{W}_3	\hat{W}_4	\hat{W}_5	\hat{W}_{sum}
SC ≤ 4	2.00	2.11	.90	.25	7.7	8.1	3.5	1.0	20.3
SC ≤ 5	1.93	1.72	.74	.24	9.6	8.6	3.7	1.2	23.2
SC ≥ 6	1.74	1.09	.47	.04	11.5	7.3	3.1	.3	22.2
								(*)	(14%)
SD ≤ 4	1.90	1.91	.95	.22	7.0	7.0	3.5	.8	18.3
SD ≤ 5	1.95	1.39	.55	.14	9.8	7.0	2.8	.7	20.2
SD ≥ 6	1.77	1.02	.39	0	11.2	6.5	2.5	0	20.1
	*								(10%)
P1 ≤ 4	1.79	1.38	.61	.07	6.7	5.1	2.3	.3	7.7
P1 ≤ 5	1.86	1.26	.45	.04	9.3	6.3	2.2	.2	8.8
P1 ≥ 6	1.69	1.21	.43	.07	10.6	7.6	2.7	.4	10.7
	*			*					(39%)
P3 ≤ 4	2.00	1.78	.52	.15	7.3	6.5	1.9	.5	9.0
P3 ≤ 5	1.83	1.15	.44	.14	9.1	5.8	2.2	.7	8.7
P3 ≥ 6	1.87	1.06	.45	.06	11.7	6.6	2.8	.3	9.8
	*		*						(14%)

TABLE XIII. \hat{N}_k and \hat{W}_k for $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 4$.

Group	\hat{N}_2	\hat{N}_3	\hat{N}_4	\hat{N}_5	\hat{W}_2	\hat{W}_3	\hat{W}_4	\hat{W}_5	\hat{W}_{sum}
SC ≤ 4	2.00	1.98	.78	.19	7.7	7.6	3.0	.7	19.1
SC ≤ 5	1.93	1.61	.66	.20	9.6	8.1	3.3	1.0	22.0
SC ≥ 6	1.72	.97	.38	.03	11.4	6.5	2.5	.2	20.6
				*					(15%)
SD ≤ 4	1.89	1.76	.83	.18	6.9	6.5	3.0	.7	17.1
SD ≤ 5	1.95	1.27	.46	.11	9.7	6.4	2.3	.5	18.9
SD ≥ 6	1.74	.89	.30	0	11.0	5.6	1.9	0	18.4
	*								(11%)
P1 ≤ 4	1.79	1.32	.55	.06	6.7	4.9	2.0	.2	7.2
P1 ≤ 5	1.85	1.16	.39	.03	9.3	5.8	2.0	.2	7.9
P1 ≥ 6	1.67	1.12	.38	.06	10.5	7.1	2.4	.4	9.8
	*			*					(36%)
P3 ≤ 4	2.00	1.72	.48	.13	7.3	6.3	1.7	.5	8.5
P3 ≤ 5	1.82	1.08	.40	.13	9.1	5.4	2.0	.6	8.0
P3 ≥ 6	1.86	1.01	.41	.05	11.6	6.3	2.6	.3	9.2
	*		*						(14%)

TABLE XIV. \hat{N}_k and \hat{W}_k for $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \infty$

Group	\hat{N}_2	\hat{N}_3	\hat{N}_4	\hat{N}_5	\hat{W}_2	\hat{W}_3	\hat{W}_4	\hat{W}_5	\hat{W}_{sum}
SC \leq 4	2.00	1.81	.67	.15	7.7	7.0	2.6	.6	17.9
SC = 5	1.91	1.48	.56	.16	9.5	7.4	2.8	.8	20.5
SC \geq 6	1.66	.84	.30	.02	11.0	5.6	2.0	.1	18.7 (15%)
				*					
SD \leq 4	1.82	1.55	.68	.14	6.7	5.7	2.5	.5	15.3
SD = 5	1.82	1.08	.36	.08	9.1	5.4	1.8	.4	16.7
SD \geq 6	1.63	.73	.22	0	10.3	4.6	1.4	0	16.3 (9%)
P1 \leq 4	1.79	1.26	.49	.05	6.7	4.7	1.8	.2	6.7
P1 = 5	1.78	1.02	.32	.02	8.9	5.1	1.6	.1	6.8
P1 \geq 6	1.59	1.00	.32	.05	10.0	6.3	2.0	.3	8.6 (28%)
				*					
P3 \leq 4	1.92	1.58	.42	.10	7.0	5.8	1.5	.4	7.7
P3 = 5	1.76	.99	.35	.11	8.8	4.9	1.7	.5	7.2
P3 \geq 6	1.80	.92	.36	.04	11.2	5.8	2.2	.2	8.2 (14%)
	*		*	*					

5.2. Summary of Data Analysis

Tables IV-XIV indicate the following general features of dendritic branching for the data sets analyzed on layer IV stellate cells from cats and layer V pyramidal cells from rats:

- a) The data fit the MBC λ model fairly well; however, the parameter λ_2 , which measures the likelihood of cutting for terminal second-order branches relative to that for bifurcating second-order branches, cannot be usefully estimated from the data. Consequently, the parameter λ_2 is treated in all subsequent analysis as an arbitrarily pre-specified parameter.
- b) The data on second-order branching is consistent with the hypothesis H_{BB} that the two sister second-order branches bifurcate independently of each other; in any event, this assumption has little effect on estimation of N_3 within the MBC λ model. Consequently, the estimation of N_4 and N_5 is carried out under the binomial branching assumption, with parameters λ_3 and λ_4 arbitrarily pre-specified.
- c) As would be expected, the second-order branching data is not consistent with the hypothesis H_{BC} that the two sister second-order branches are cut independently of each other.
- d) The main conclusion of [2] is that the dendritic branching pattern tends to compensate for disparate numbers of trees per neuron in such a way as to maintain constancy of

W_{sum} , the mean number of higher-order branches per neuron. From Tables VIII-XIV, this phenomenon is seen to hold for the SC, SD and P3 data sets over a wide range of possible values of the unknown parameters λ_k . The P1 data set shows a less pronounced tendency toward this same phenomenon.

APPENDIX

In this appendix we present the equations for determining the maximum likelihood estimates of the second-order branching parameters (p_{ij}) and cutting parameters ($\overline{b_2}$ and $\overline{b_2^2}$) under the MBC λ model through differentiation of the log likelihood function. See Equations (3.6) and (3.11) for the definition of the likelihood function L . To simplify the typography of the formulas, the subscripts will be omitted from λ_2 , $\overline{b_2}$ and $\overline{b_2^2}$.

The branching parameters p_{ij} are estimated from the equations

$$\begin{aligned} \frac{\partial \log L}{\partial p_{12}} &= \frac{k_2(2\lambda - 2\lambda^2)}{p_{11}\lambda^2 + 2p_{12}\lambda + p_{22}} + \frac{n_1(1 - 2\lambda)}{\lambda p_{11} + p_{12}} \\ &+ \frac{n_2\lambda}{\lambda p_{12} + p_{22}} - \frac{2m_{11}}{p_{11}} + \frac{m_{12}}{p_{12}} = 0 \end{aligned} \quad (A1)$$

$$\begin{aligned} \frac{\partial \log L}{\partial p_{22}} &= \frac{k_2(1 - \lambda^2)}{p_{11}\lambda^2 + 2p_{12}\lambda + p_{22}} - \frac{\lambda n_1}{\lambda p_{11} + p_{12}} + \frac{n_2}{\lambda p_{12} + p_{22}} \\ &- \frac{m_{11}}{p_{11}} + \frac{m_{22}}{p_{22}} = 0 \end{aligned} \quad (A2)$$

and
$$p_{11} = 1 - 2p_{12} - p_{22}. \quad (A3)$$

The maximum likelihood estimates \hat{p}_{11} , \hat{p}_{12} , \hat{p}_{22} are the simultaneous solutions of A1, A2 and A3.

The cutting parameters $\overline{b_2}$ and $\overline{b_2^2}$ are estimated from the equations

$$\begin{aligned}
\frac{\partial \log L}{\partial \bar{b}} &= \frac{n_1}{\bar{b} - \lambda \bar{b}^2} + \frac{n_2}{\bar{b} - \bar{b}^2} \\
&- \frac{2\lambda m_{11}}{1 - 2\lambda \bar{b} + \lambda^2 \bar{b}^2} - \frac{(\lambda + 1)m_{12}}{1 - (\lambda + 1)\bar{b} + \lambda \bar{b}^2} \\
&- \frac{2m_{22}}{1 - 2\bar{b} + \bar{b}^2} = 0
\end{aligned} \tag{A4}$$

$$\begin{aligned}
\frac{\partial \log L}{\partial \bar{b}^2} &= \frac{k_2}{\bar{b}^2} - \frac{\lambda n_1}{\bar{b} - \lambda \bar{b}^2} - \frac{n_2}{\bar{b} - \bar{b}^2} \\
&+ \frac{\lambda^2 m_{11}}{1 - 2\lambda \bar{b} + \lambda^2 \bar{b}^2} + \frac{\lambda m_{12}}{1 - (\lambda + 1)\bar{b} + \lambda \bar{b}^2} + \frac{m_{22}}{1 - 2\bar{b} + \bar{b}^2} \\
&= 0
\end{aligned} \tag{A5}$$

Computational difficulties caused by large or infinite values of λ can be avoided by using the natural symmetry of the problem: using as input $1/\lambda$ instead of λ , interchanging n_1 and n_2 , interchanging m_{11} and m_{22} , then on output reading \bar{b} as \bar{t} , \bar{b}^2 as \bar{t}^2 , p_{11} as p_{22} , and p_{22} as p_{11} .

REFERENCES

- [1] McCabe, G. P., Jr., and Samuels, M. L. (1977) Random censoring and dendritic trees. Biometrics 33, 69-84.
- [2] Samuels, M. L., Mittenthal, J. E., McCabe, G. P., Jr., and Coleman, P. D. (1977) Complexity of branching in dendritic trees: dependence on number of trees per cell and effects of branch loss during sectioning. Journal of Anatomy 124, 3, 701-715.
- [3] Smit, G. J., Dylings, H. B. M. and Veldmaat-Wansink, L. (1972) The branching pattern in dendrites of cortical neurons. Acta Morphologica Neerlando-Scandinavica 9, 253-274.
- [4] Wilks, Samuel S. (1962) Mathematical Statistics. New York: John Wiley & Sons.