Journal of Neuroscience Methods, 14 (1985) 137–142
Elsevier

NSM 00489

A simple statistical test for the vertex ratio using Monte Carlo simulation

Ronald W.H. Verwer, Jaap van Pelt and Harry B.M. Uylings

Netherlands Institute for Brain Research, Amsterdam (The Netherlands)

(Received January 9th, 1985)
(Revised March 22nd, 1985)
(Accepted March 25th, 1985)

Key words: neuronal branching pattern – topological property – vertex analysis – Monte Carlo test – growth hypothesis

The vertex ratio is the crucial quantity in vertex analysis, which is a method to characterize the mode of growth of neuronal tree structures (i.e. dendrites and axons). In this report we propose the use of the Monte Carlo test to calculate a level of significance for the vertex ratio. As a result the vertex ratio can be used to analyse neuronal trees with respect to a range of growth hypotheses, including terminal and segmental growth.

Introduction

Recently some progress has been achieved in characterizing and testing the topological properties of neuronal tree structures (cf. Berry and Flinn, 1984; Van Pelt and Verwer, 1983; Verwer and Van Pelt, 1983). One of the methods to characterize the mode of growth that might be responsible for the formation of the observed neuronal trees is called vertex analysis. The crucial quantity in vertex analysis, the vertex ratio, is defined as the ratio of the number of primary and the number of secondary nodal vertices (Berry and Flinn, 1984). A primary nodal vertex is a bifurcation point with two terminal segments and a secondary nodal vertex is a bifurcation point with one terminal and one intermediate segment. It was shown that the vertex ratio is 1.0 for terminal growth and approaches asymptotically 0.5 in large trees for segmental growth (Berry and Flinn, 1984; Verwer and Van Pelt, 1985). However, the vertex ratio can only be used as an indication of the mode of growth since there is no level of significance associated with it. In this report we propose a statistical procedure for the vertex ratio, which previously has been used by Besag and Diggle (1977) to test spatial pattern. This kind of testing, called Monte Carlo

Correspondence: R.W.H. Verwer, Netherlands Institute for Brain Research, Meibergdreef 33, 1105 AZ Amsterdam, The Netherlands.

test, was originally introduced by Barnard (1963). With this procedure it is possible to use the vertex ratio with a whole range of growth hypotheses, which include terminal and segmental growth as particular cases. An analytical expression for these growth models has been derived by Van Pelt and Verwer (1985) and is given below.

The Monte Carlo test

The approach can be summarized as follows: Suppose a set of m trees of degree n (i.e. trees with n terminal segments) was observed and the vertex ratio has been calculated subsequently. To verify whether the formation of the observed trees is in agreement with a certain hypothetical growth mode one simulates for the considered growth hypothesis e.g. 99 times a sample of m trees of degree n and computes the 99 corresponding vertex ratios. Next, the vertex-ratio values are ranked and the rank of the observed value is used to find the significance level. This is illustrated in the example (see below). The reader may also refer to the section on sampled randomization tests in Sokal and Rohlf (1981). A disadvantage connected with this kind of testing is that the critical region is not well defined and that different investigators may sometimes arrive at different conclusions with the same set of data (Besag and Diggle, 1977). This vagueness of the critical level is not very serious when the number of simulated values that lead to rejection is not less than 5 (Marriott, 1979). This means that, as was already noted by Besag and Diggle (1977), in most cases 99 simulations will be satisfactory (e.g. for 100 ranks including the observed value a critical level $\alpha = 0.05$ corresponds to 5 ranks). It may be convenient in two-sided tests to use 199 simulations in which case rejection is concluded if the observed vertex ratio is among the 5 largest or among the 5 smallest values. One may consider to increase the number of simulations when the rank of the observed vertex ratio is close to the critical level.

Example

A sample of 6th-degree basal dendritic trees of 14-day-old pyramidal cells from layer II/III in the visual cortex of the rat is used to illustrate the procedure. The 6 possible tree types of degree 6 are displayed in Table I together with their observed frequencies. It was shown (Verwer and Van Pelt, 1985) that the vertex-ratio estimate can be determined from the mean number of 2nd-degree subtrees present in the observed nth-degree trees $E\{f(2|n)\}$, according to relation (1). Here, $f(2|n)$ denotes the number of 2nd-degree subtrees in a tree of degree n and $E\{\cdot\}$ denotes the expectation or mean value. The total number of 2nd-degree subtrees appeared to be 63 occurring in the 33 observed 6th-degree trees and thus results in a mean number:
$E\{f(2|6)\}_{obs} = 1.91$
which in turn gives as a value for the vertex-ratio estimate:

$$R_{obs} = \frac{E\{f(2|n)\}_{obs}}{n - 2 \cdot E\{f(2|n)\}_{obs}} = \frac{1.91}{6 - 3.82} = 0.875 \qquad (1)$$

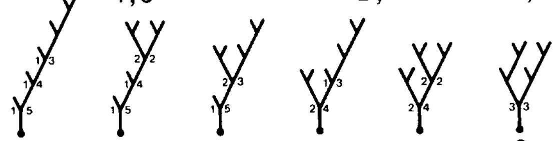
(cf. Verwer and Van Pelt, 1985). For reasons of space we have in this example only

TABLE I

A SAMPLE OF 33 6TH-DEGREE BASAL DENDRITIC TREES OF PYRAMIDAL NEURONS OF LAYER II/III IN THE VISUAL CORTEX OF THE RAT

The number of times each tree type was observed is shown together with the amount of 2nd-degree subtrees they contribute. The numbers above the types indicate how the terminal segments are partitioned over the first-order bifurcation point and serve as a means for classifications (cf. Van Pelt and Verwer, 1983). The same holds for the numbers on both sides of each bifurcation point that is relevant for topological analysis. Thus, the trees may be classified as 6(1,5), 6(2,4) and 6(3,3), the 5th-degree subtrees as 5(1,4) and 5(2,3) and the 4th-degree subtrees as 4(1,3) and 4(2,2).

Type: 1,5 2,4 3,3



Rank:	1	2	3	4	5	6	
Observed frequency:	6	–	8	12	3	4	*Total* 33
Contribution of 2nd-degree subtrees:	6	–	16	24	9	8	63

simulated 19 vertex-ratio values (instead of the advised 99) under the hypothesis of terminal growth and ranked them including the observed value (between brackets):

0.769	0.914	1.000	1.047
0.803	0.956	1.000	1.097
0.838	1.000	1.000	1.150
0.875	1.000	1.000	1.207
(0.875)	1.000	1.047	1.268

There are 5 values out of 20 in this list that are equal or smaller than 0.875, therefore we accept terminal growth at a level of significance of 25%. The procedure consists of determining a reasonably accurate frequency distribution of the estimator (vertex ratio) for the particular sample of trees and the specific growth hypothesis, because the actual probability distribution is unknown. As in normal statistical practice we check whether the observed estimator is unusually small or large with respect to the values it is expected to assume under the hypothesis. Two aspects are noteworthy. We have used the fact that the observation is lower than the expected vertex ratio, which under terminal growth is 1.000, and applied a one-sided test. And since the distribution of the vertex ratio is discrete, the level of significance includes the simulated values that are equal to 0.875 (cf. Besag and Diggle, 1977).

For the simulation of the vertex-ratio values we did not generate the distribution of the types shown in Table I but instead used the distributions of the possible

TABLE II

LIST OF THE SIMULATED 'OBSERVATIONS' THAT CORRESPOND TO THE VERTEX-RATIO VALUE 1.150 (SEE DESCRIPTION IN THE TEXT)

With respect to relevant topological information the 'data' generated in this way are equivalent to a sample of 33 complete 6th-degree trees. The number of 2nd degree subtrees is determined as follows: A subtree of degree 3 always contains one 2nd-degree subtree and thus a subtree with partition (1,3) and a subtree with partition (3,3) contribute one and two 2nd-degree subtrees respectively. A subtree type with partition (2,2) clearly contributes two 2nd-degree subtrees. A subtree with partition (1,4) does not contribute any 2nd-degree subtree since the subtree of degree 1 cannot contain a 2nd-degree subtree and the 2nd-degree subtrees in the subtree of degree 4 are already counted.

Subtree partition class	Number of observations generated	Corresponding number of 2nd-degree subtrees
4(1,3)	11	11
4(2,2)	10	20
5(1,4)	7	–
5(2,3)	7	14
6(1,5)	14	–
6(2,4)	14	14
6(3,3)	5	10
		Total 69

partitions into subtree pairs. This has considerable advantages when larger trees or mixed populations have been observed. The hypothetical probability distributions of the partitions into subtree pairs can be calculated (Van Pelt and Verwer, 1985) as follows:

$$p(r,n-r;Q)=\frac{2+Q(n-4)}{(n-1-Q)} \qquad \text{if } r=1$$

$$p(r,n-r;Q)=\prod_{i=1}^{r-1}\left(\frac{i-Q}{i}\right)\cdot\prod_{i=n-r}^{n-1}\left(\frac{i}{i-Q}\right) \qquad (2)$$

$$\cdot\left[1+Q\left\{\frac{n(n-1)}{2r(n-r)}-2\right\}\right]\cdot\frac{2^{1-\delta_{r,n-r}}}{(n-1)} \qquad \text{if } 1<r\leqslant n-r$$

Here δ is the Kronecker delta and equals 0 if $r\neq n-r$ and is 1 if $r=n-r$. Further, n denotes the degree of the trees, r is the degree of the smalles first-order subtree and Q ($0\leqslant Q\leqslant 1$) is a parameter defining the growth hypothesis ($Q=0$ for terminal growth and $Q=0.5$ for segmental growth). Thus, to simulate a vertex-ratio value in the above list for terminal growth, the probabilities $p(1,5;Q=0)$, $p(2,4;Q=0)$ and $p(3,3;Q=0)$ were calculated for the classes 6(1,5), 6(2,4) and 6(3,3) respectively (see also Table I) and 33 times a random number W between 0 and 1 was generated. If W was less than or equal to $p(1,5;Q=0)$ one 'observation' was added to the class 6(1,5), when $p(1,5;Q=0)<W\leqslant p(1,5;Q=0)+p(2,4;Q=0)$ one 'observation' was added to class 6(2,4) and when $p(1,5;Q=0)+p(2,4;Q=0)<W\leqslant 1$ one 'observa-

tion' was added to class 6(3,3). Suppose that in this way 14 'observations' would have been generated for class 6(1,5), then we would calculate the probabilities $p(1,4;Q=0)$ and $p(2,3;Q=0)$ and generate 14 'observations' for the classes 5(1,4) and 5(2,3) as described above. If 7 'observations' are allocated to class 5(1,4) and if 14 'observations' were obtained for class 6(2,4) we proceed to generate 21 'observations' for classes 4(1,3) and 4(2,2) after we calculated the probabilities $p(1,3;Q=0)$ and $p(2,2;Q=0)$. If class 4(1,3) would have received 11 'observations' and therefore class 4(2,2) received 10 we have generated the Monte Carlo data-set shown in Table II. A number of 69 2nd-degree subtrees in 33 simulated trees of degree 6 yields a value of the vertex-ratio of 1.150 (see list of generated vertex values above).

Evaluation

We have performed 10 one-sided Monte Carlo tests of 99 simulations each for both terminal and segmental growth. The expectation of the vertex-ratio estimate under segmental growth takes the value

$$R=\left\{\frac{2(2n-3)}{(n-1)}-2\right\}^{-1}=\left\{\frac{18}{5}-2\right\}^{-1}=0.625 \tag{3}$$

for 6th-degree trees (cf. Verwer and Van Pelt, 1985). Consequently the level of significance for segmental growth is the percentage of values larger than or equal to 0.875 and for terminal growth it is the percentage of values smaller than or equal to 0.875. The obtained levels of significance were: 0.02, 0.02, 0.02, 0.04, 0.03, 0.04, 0.05, 0.04, 0.02 and 0.01 for segmental growth hypothesis and 0.18, 0.24, 0.22, 0.21, 0.21, 0.24, 0.30, 0.21, 0.20 and 0.18 for terminal growth hypothesis. These results might have been obtained by ten different investigators each performing the Monte Carlo test once for both hypotheses. This illustrates the effect of 'blurring' of the critical level and the corresponding loss of power (Hope, 1968; Marriott, 1979). However, the conclusions are evident for each test: terminal growth is accepted and segmental growth is rejected. As was suggested above if a level of significance of 0.06 or 0.05 is obtained one may decide to increase the number of simulations. To compare the performance of this testing procedure we have analysed the same data set with the subtree partition analysis (SPA) (Verwer and Van Pelt, 1983). The level of significance for terminal growth was 0.33–0.36, whereas it was 0.01 for segmental growth leading to the same conclusions as obtained above. Using the Monte Carlo procedure and formulae (1) and (2) it is possible to extend the application of the vertex ratio to growth hypotheses other than terminal and segmental growth. It thus appears that the Monte Carlo test enables one to use the vertex ratio for hypothesis testing, although it must be kept in mind that the vertex ratio only uses a small part of the available topological information. Finally, the Monte Carlo test can be performed equally well with the mean number of 2nd-degree subtrees, since the vertex ratio is uniquely determined by the number of 2nd-degree subtrees.

Acknowledgements

We are grateful to G. van der Meulen and H. Stoffels for their help in the presentation of the tree types.

References

Barnard, G.A. (1963) Discussion of Professor Bartlett's paper, J. roy. statist. Soc. B, 25: 294.

Berry, M. and Flinn, R. (1984) Vertex analysis of Purkinje cell dendritic trees in the cerebellum of the rat, Proc. roy. Soc. B, 221: 321–348.

Besag, J. and Diggle, P.J. (1977) Simple Monte Carlo tests for spatial pattern, Appl. Statist., 26: 327–333.

Hope, A.C.A. (1968) A simplified Monte Carlo significance test procedure, J. roy. statist. Soc. B, 30: 582–598.

Marriott, F.H.C. (1979) Barnard's Monte Carlo tests: how many simulations?, Appl. Statist., 28: 75–77.

Sokal, R.R. and Rohlf, F.J. (1981) Biometry: The Principles and Practice of Statistics in Biological Research, 2nd edn., Freeman, San Fransisco, pp. 791–795.

Van Pelt, J. and Verwer, R.W.H. (1983) The exact probabilities of branching patterns under terminal and segmental growth hypothesis, Bull. math. Biol., 45: 269–285.

Van Pelt, J. and Verwer, R.W.H. (1985) Growth models (including terminal and segmental branching) for topological binary trees, Bull, math. Biol., 47: in press.

Verwer, R.W.H. and Van Pelt, J. (1983) A new method for the topological analysis of neuronal tree structures, J. Neurosci. Meth., 8: 335–351.

Verwer, R.W.H. and Van Pelt, J. (1985) Topological analysis of binary tree structures when occasional multifurcations occur, Bull. math. Biol., 47: 305–316.