RESEARCH REPORTS

Department of Statistics

University of Leeds

Leeds LS2 9JT

SOME STATISTICAL INFERENCE PROBLEMS IN KRIGING I :

NUMERICAL APPLICATIONS

by

K.V. Mardia and C.A. Gill

No.5 August, 1980

Some statistical inference problems in kriging I : numerical applications.

K.V. Mardia and Christine A. Gill

SUMMARY

By using examples, the paper shows that the quantitative study of anisotropy is important. At present, only qualitative methods are available to assess isotropy. We give a test of isotropy as well as estimators of the parameters for a doubly-geometric process. Further, we provide a test based on a variance-stabilising transformation. Also, some log plots are described. Various numerical examples are given. The theory of the results used in this paper is given in Mardia (1980).

1. Isotropy hypothesis.

The basic tool for kriging is the variogram and in this paper we will concentrate mainly on the hypothesis of isotropy. Let $\{z(\underset{\sim}{x})\}$ be a stationary stochastic process. Let $\underset{\sim}{h}$ be the vector of distances. Then the semi-variogram is

$$\gamma(\underset{\sim}{h}) = \frac{E\{z(\underset{\sim}{x}+\underset{\sim}{h}) - z(\underset{\sim}{x})\}^2}{2} .$$

Its sample counterpart is given by

$$v(\underset{\sim}{h}) = \frac{\sum_{i=1}^{n(h)} \{z(\underset{\sim}{x}_i+\underset{\sim}{h}) - z(\underset{\sim}{x}_i)\}^2}{2n(\underset{\sim}{h})} ,$$

where $\underset{\sim}{h}$ is the distance or lag between pairs of points and $n(\underset{\sim}{h})$ is the number of pairs of observations at lag $\underset{\sim}{h}$.

It is usually assumed that $\gamma(\underset{\sim}{h})$ is isotropic, ie.

$$\gamma(\underset{\sim}{h}) = \gamma_1(|\underset{\sim}{h}|) = \gamma_1(h), \quad \text{say.}$$

Thus, it only depends on the length h of the vector $\underset{\sim}{h}$. We will consider isotropy in the restricted sense given by the definitions in §2. The isotropic variogram is not only theoretically simple but also appears in practice. Further, the variogram has an advantage over the covariogram and correlogram because it has three natural parameters, nugget effect, range and sill. Also, for irregular points, the spectrum is not a natural tool.

It is sometimes believed that the study of anisotropy is not important but the following quotation from Krige, (1978, p.35), is worth remembering.

"The attention given in this section to anisotropy is justified by the fact that, where it is significant, *its recognition, proper definition and introduction into ore valuation procedures will always yield* *significant improvements in efficiency."*

The italics are as in the original text.

However, the most common method to examine anisotropy is a qualitative (or ad-hoc) graphical procedure. The variograms are plotted in four directions, N/S, E/W, NE/SW, NW/SE and anisotropy is assessed by looking at their structure. For example, Fig. 1 shows the variograms from Krige (1976) for an African gold mine. Krige concludes that it shows significant anisotropy. But in general, this judgement requires a considerable amount of experience.

Insert Fig.1.

It is not always possible to make a clear cut judgement from diagrams. Figs.2 and 3 give variograms in only two directions for two sets of simulated data from a geometrical model (see §2). It has been found in conferences that, in general, one cannot state clearly which is isotropic and which is not. One impression is given by looking at the large lags and a different impression is given by the small lags. These figures are discussed further in §2.

Insert Figs.2,3.

We therefore need a test of isotropy with a computing algorithm which can automatically tell us how far away we are from isotropy. In fact, Matheron (1961) was the first to address this question which arose in connection with a bauxite deposit data. As far as we can gather, he suggested a χ^2 test of equality of fit of the E/W and N/S distributions for $h = 1$. This will detect anisotropy for very small lags but the data contains more information. His variograms indicate that the data is isotropic, (see Fig. 4).

Insert Fig.4.

The mathematical difficulty arises because we have dependence of each element of $\underset{\sim}{V}_1$ in the N/S direction and of $\underset{\sim}{V}_2$ in the E/W direction as well as their mutual dependence.

2. The doubly geometric process and related models.

One method to tackle this problem is to use suitable stochastic processes on a regular lattice in 2 dimensions. One of the simplest models in 2-D is

$$z_{ij} = \lambda z_{i-1,j} + \nu z_{i,j-1} - \lambda\nu z_{i-1,j-1} + \varepsilon_{ij}, \quad |\lambda| < 1, \ |\nu| < 1,$$

where ε_{ij} are independent $N(0,\sigma^2)$ variates, independent of z_{ij}. It

isert
g.5.

models spatial dependence of z_{ij} in three directions, (see Fig. 5). This leads to

$$\gamma(\underset{\sim}{h}) = \sigma_z^2(1-\lambda^{h_1}\nu^{h_2}) \quad , \tag{2.1}$$

where

$$\sigma_z^2 = \text{var}(z_{ij}).$$

The theoretical variograms for this process are shown in Fig.6 for our previous examples where $\lambda = \nu = 0.5$, (Fig. 2), and $\lambda = 0.6$, $\nu = 0.4$, (Fig.3). For $\lambda = \nu$, we have isotropy but for $\lambda \neq \nu$, anisotropy. Note

isert
g.6.

that both these examples are laterally symmetric, ie. invariant under reflection of the N/S and E/W directions, but only the example of $\lambda = \nu$ is symmetrical in all the four direction N/S, E/W, NE/SW, NW/SE and hence "isotropic" in this sense.

A likelihood ratio test can be written down after some lengthy mathematical manipulation. It can also be shown that, asymptotically,

$$\{\sin^{-1}\lambda - \sin^{-1}\nu\}\sqrt{N} \sim N(0,2).$$

We can use the likelihood ratio test to examine

$$H_o: \lambda = \nu$$

versus

$$H_1: \lambda \neq \nu$$

with our data which has been simulated on a 15×15 grid for the doubly-

Insert Table 1.

geometric process, (see Table 1). Maximum likelihood estimates of λ, ν and σ^2 for the simulated data are as follows: for the data with $\lambda = \nu = 0.5$, $\hat{\lambda} = 0.54$, $\hat{\nu} = 0.59$, $\hat{\sigma}^2 = 1.00$, while for the non-isotropic data with $\lambda = 0.6, \nu = 0.4$, we have $\hat{\lambda} = 0.62$, $\hat{\nu} = 0.43$, $\hat{\sigma}^2 = 1.00$. We correctly accept the null hypothesis for the data with $\lambda = \nu$ and reject it when $\lambda \neq \nu$. Hence we can apply this test which asymptotically depends on the variograms of lag 1 in directions N/S, E/W, NE/SW, NW/SE, like Matheron's test. In fact, for large m,n, and small λ, ν and γ we have approximately

$$- 2 \log \Lambda \simeq \frac{N(v_{11}-v_{21})^2}{\{2M_{oo}(v^*_{11}+v^*_{21}-2M_{oo})\}} \simeq \chi^2_1 ,$$

where $M_{oo} = \Sigma\Sigma z^2_{ij}$, $v_{11}, v_{21}, v^*_{11}, v^*_{21}$ are the variograms along 0^o, 90^o, 45^o and 135^o respectively.

The model is restictive because it does not use variograms of other lags. However, it is simple and it is the most efficient method if the model is correct. Journel and Huijbregts (1978, p. 233-4) give evidence that a nested spherical scheme can be approximated by (2.1) for $\lambda = \nu$. Hence, it should have a wider applicability.

In fact, for a 15×15 isotropic data set, simulated from the spherical scheme with range $= 6$, sill $= 0.2468$ and nugget effect $= 0.05$, we find that $-2 \log \Lambda = 2.67$ and we can again accept the null hypothesis.

It should also be noted that the behaviour of (2.1) is similar to that of $\gamma(\underset{\sim}{h}) = \sigma^2_z(1-\lambda^{h_1^2}\nu^{h_2^2})$ for h_1 and h_2 small.

3. The variance-stabilising transformation.

Let $\underset{\sim}{v}'_1 = (v_{11}, v_{21}, \ldots, v_{1p})$ and $\underset{\sim}{v}'_2 = (v_{21}, v_{22}, \ldots, v_{2p})$. Under certain assumptions, $u_i = \log \dfrac{v_{1i}}{v_{2i}}$, $i = 1,2,\ldots,p$, are related in a linear Markov form. Note that the variances of the u_i are independent of scale, (ie, of $\mathrm{var}\{z(\underset{\sim}{x})\}$). For small β, an asymptotic test is then,

$$S = \frac{N}{2(1-\hat{\beta}^2)} \sum_{i=2}^{p} (u_i - \hat{\beta} u_{i-1})^2 \sim \chi^2_{p-2} \qquad (3.1)$$

where $\hat{\beta}$ is the least-squares estimate. Choice of p should be made with regard to the range of the data. This test does not take into account the edge effects, small sample sizes, etc. Various modifications taking these into consideration are in progress.

Applying this test to the Matheron data, to Krige's data, (after interpolation from Fig. 1), and to the simulated spherical data described in §2, we have the following result, (see Table 2). We certainly accept that Krige's data is anisotropic and we can accept that both the Matheron data and the simulated spherical data are isotropic at the 1% level.

Insert Table 2.

Further, note that plotting u_i provides a more interpretable picture than just plotting (v_{1i}, v_{2i}). In the Matheron example, (see Fig. 7), we have negative as well as positive values and symmetry about zero. For Krige's data, (see Fig. 8), all the values are positive and away from zero, i.e, it is asymmetric.

Insert Fig. 7.

Insert Fig. 8.

4. <u>The direction of anisotropy</u>.

We have concentrated on only two directions which are orthogonal. In practice, the preferential direction of anisotropy is generally known a priori, e.g, there is more variation in the vertical direction of a mine than in the horizontal direction. If this is not the case, then one can take λ corresponding to the major axis and ν corresponding to the minor axis of the ellipse of ranges, (see Fig. 9). However, the continuous version of the doubly-geometric process requires some precaution because of its peculiar behaviour at $\theta = 0^\circ, 45^\circ, 90^\circ$ and 135°. Various implications are examined in Mardia (1980).

Insert Fig. 9.

Conclusions

This is just the beginning of a very difficult topic. Various procedures have also been developed for estimating $\gamma(\underset{\sim}{h})$ under different models such as the spherical scheme. The theory behind the above paper as well as the additional topics is given in Mardia (1980).

Acknowledgements

We would like to thank Andrew Baczkowski, David Cox, Rob Edwards and John Kent for their valuable help and comments.

References

JOURNEL, A.G. and HUIJBREGTS, Ch.J.(1978). Mining Geostatistics. Academic Press.

KRIGE, D.G. (1976). Some basic considerations in the application of geostatistics to gold ore valuation. J.S.African Inst.Mining and Metall. 76, pp.383-391.

KRIGE, D.G. (1978). Log normal-de Wijsian geostatistics for ore valuation. S.African Inst. Mining and Metall.

MARDIA, K.V. (1980). Some statistical inference problems in Kriging II: Theory, Int.Geol.Congr. 26 (to appear).

MATHERON, G. (1961). Precision of exploring a statistical formation by boreholes with rigid spacing - application to a bauxite deposit. Int.Symp. Mining Research. University of Missouri I, 407-422.

AUTHORS

MARDIA, K. V. and GILL, C.A. Department of Statistics, University of Leeds, Leeds LS2 9JT.

List of figures

Fig. 1. Directional variograms for Subsection 302 of the Hartebeesfontein Mine, taken from Krige, (1976).

Fig. 2. Variograms in N/S and E/W directions for simulated data from a doubly-geometric model with $\lambda = \nu$.

Fig. 3. Variograms in N/S and E/W directions for simulated data from a doubly-geometric model with $\lambda \neq \nu$.

Fig. 4. Variograms in N/S and E/W directions for bauxite deposit data of Matheron (1961).

Fig. 5. Spatial dependence of z_{ij} in three directions.

Fig. 6. Theoretical variograms for the simulated doubly-geometric data.

Fig. 7. Log-plots for the bauxite deposit data of Matheron (1961).

Fig. 8. Log-plots for the gold data of Krige (1976).

Fig. 9. Ellipse of ranges with major axis $= \lambda$, minor axis $= \nu$.

List of Tables

Table 1. Likelihood ratio test of isotropy for the simulated doubly-geometric data.

Table 2. Application of S-test of isotropy.

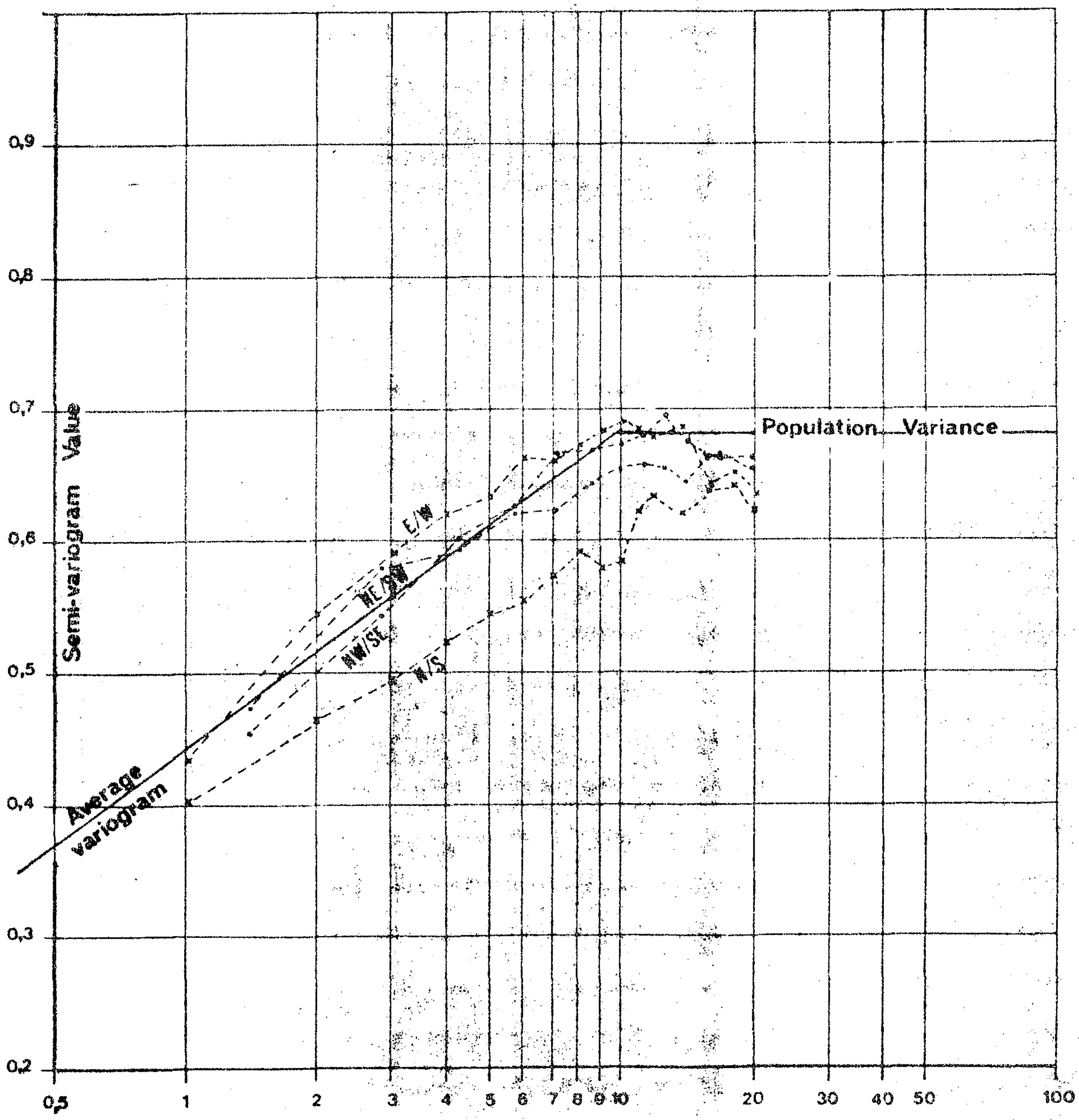
TABLE 1

Doubly Geometric Model	L.R.T.	$\chi^2_{5\%}$	$\chi^2_{1\%}$
$\lambda = \nu = 0.5$	0.54	3.84	6.63
$\lambda = 0.6,\ \nu = 0.4$	10.54	3.84	6.63

TABLE 2

Data	N	S	$\chi^2_{5\%}$	$\chi^2_{1\%}$	d.f.
Bauxite data(Matheron, 1961)	300	17.5	15.5	20.0	8
Gold data(Krige, 1976)	400	58.7	19.7	24.7	11
Simulated spherical scheme	225	12.8	9.49	13.3	4

TOP


Semi-variogram Value
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,5
1
2
3
4
5
6
7
8
9
10
20
30
40
50
100
Population Variance
E/W
NE/SW
NW/SE
N/S
Average variogram

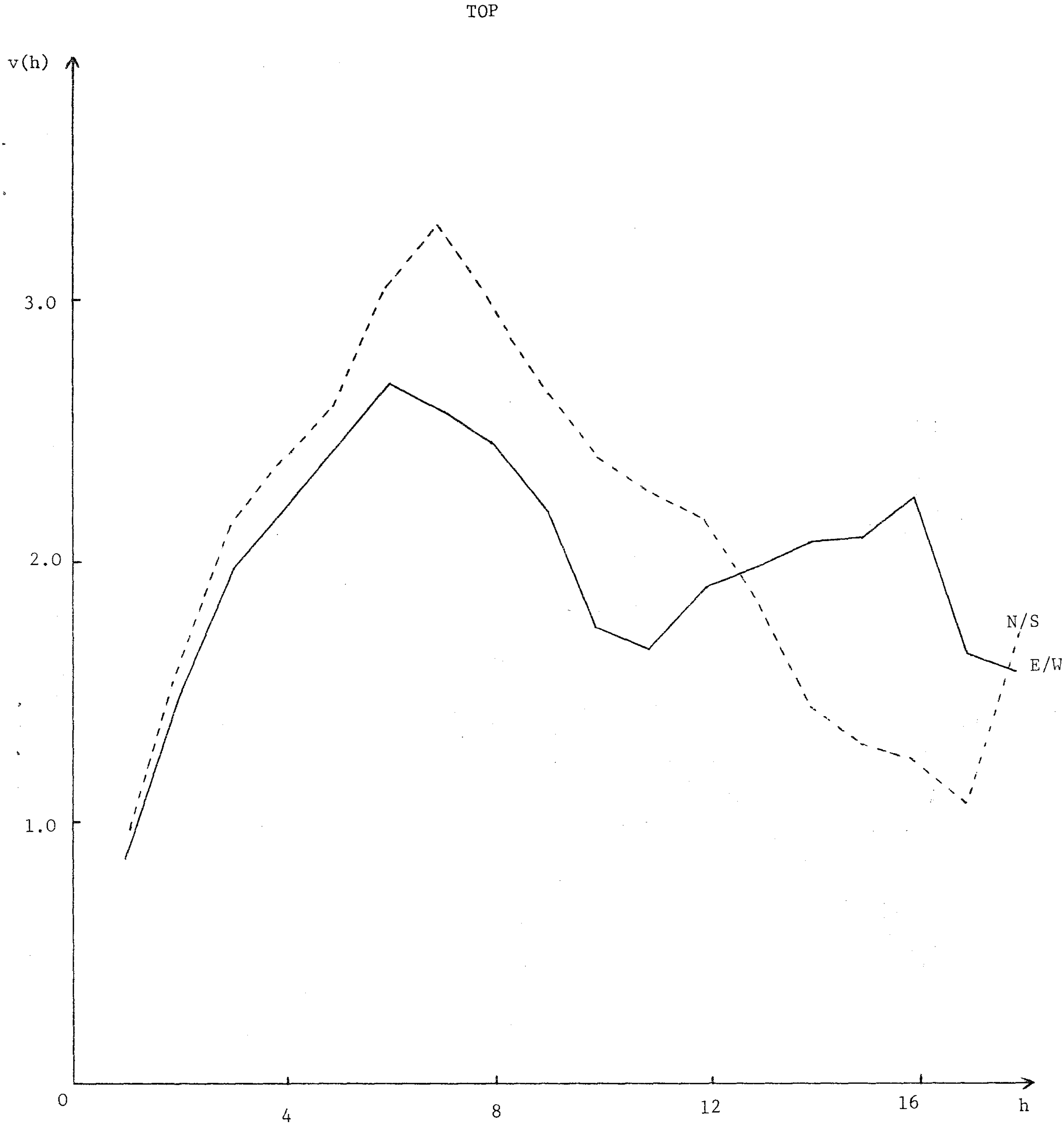


FIG.2.

TOP

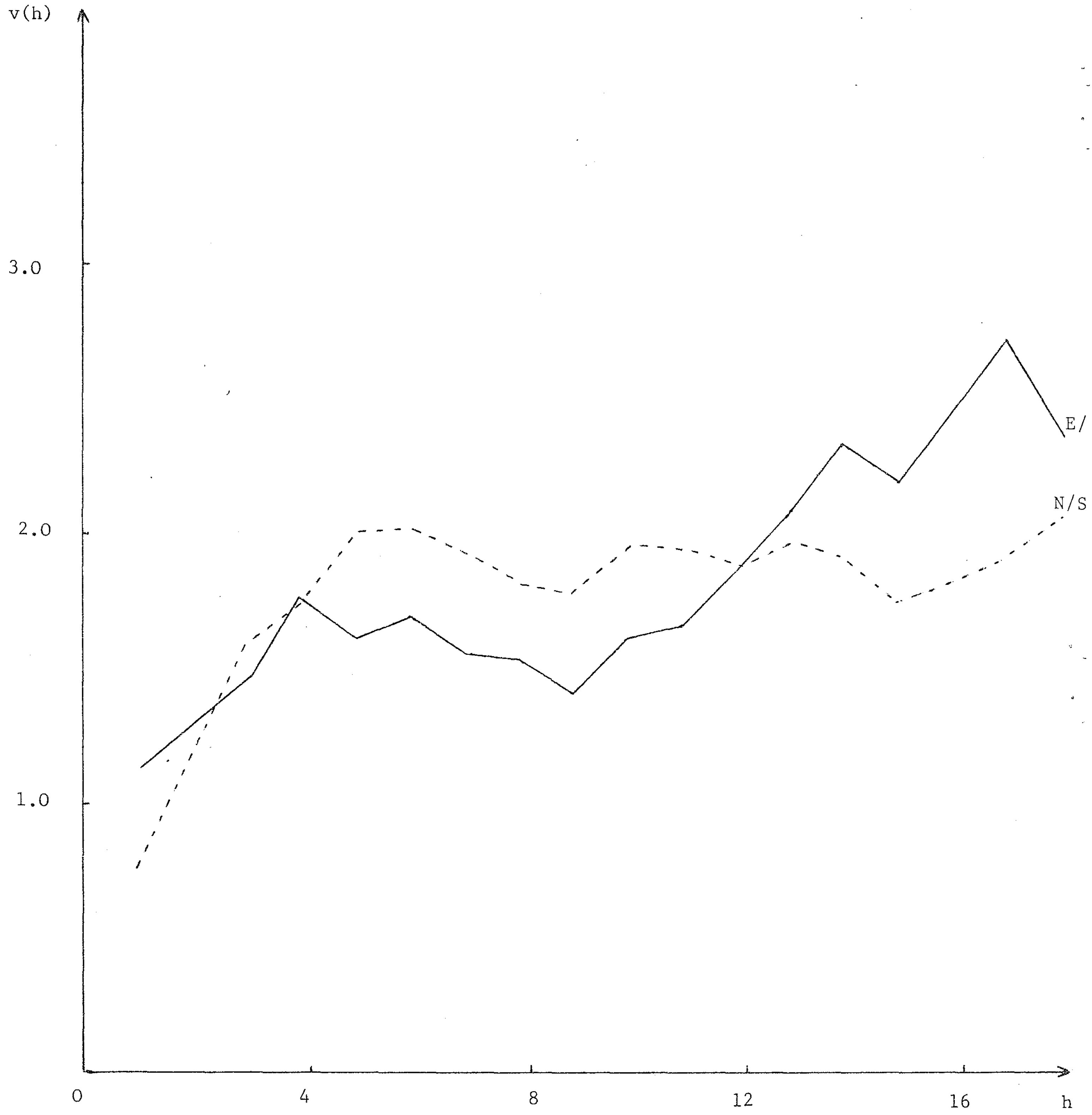


FIG.3.

TOP

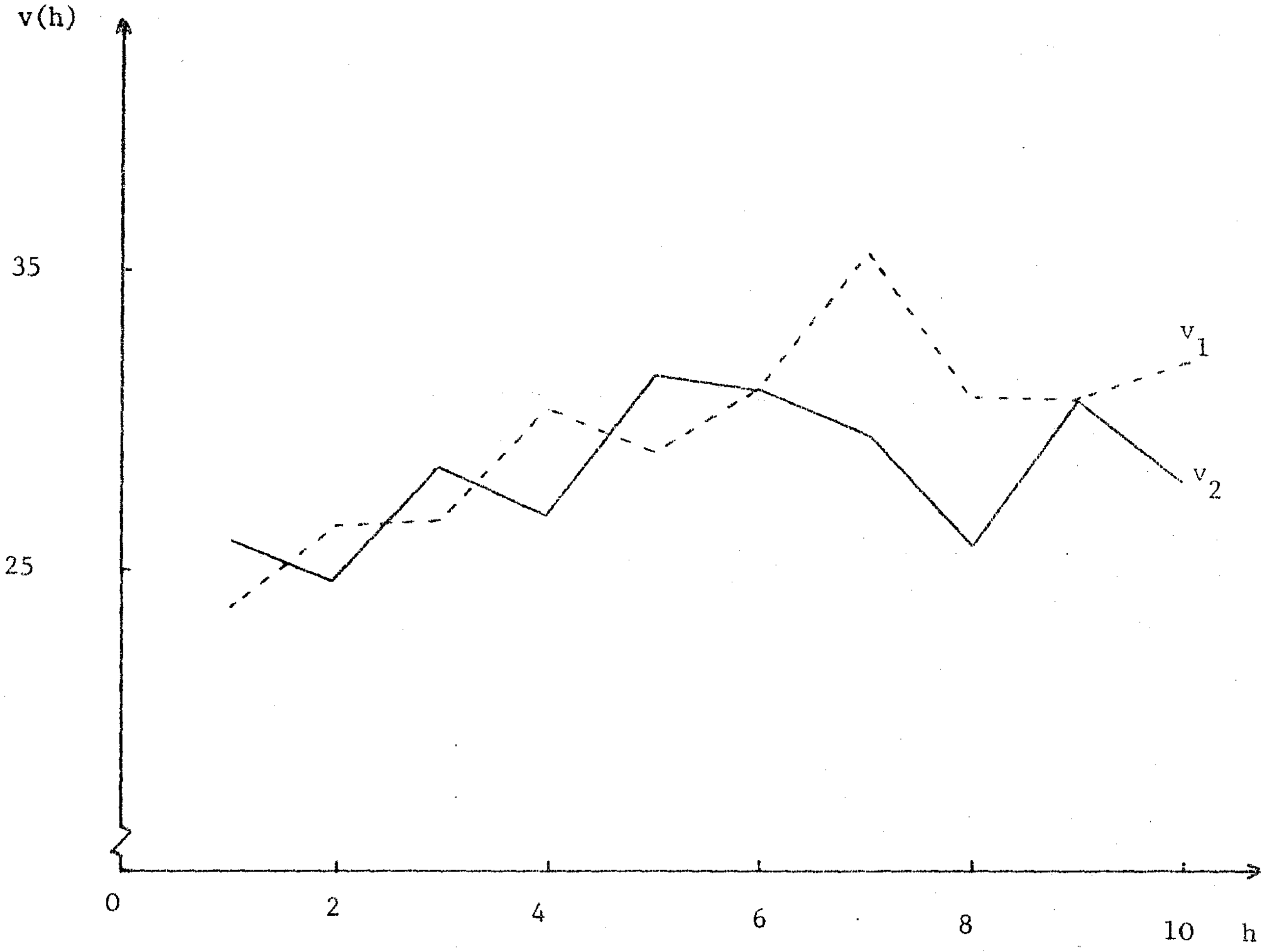


FIG.4.

TOP

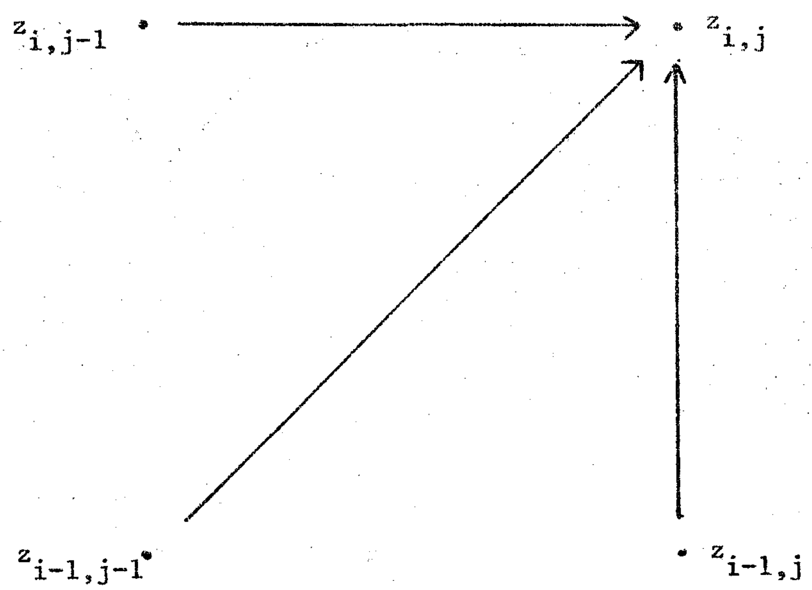


FIG.5.

K.V. Mardia & C.A. Gill

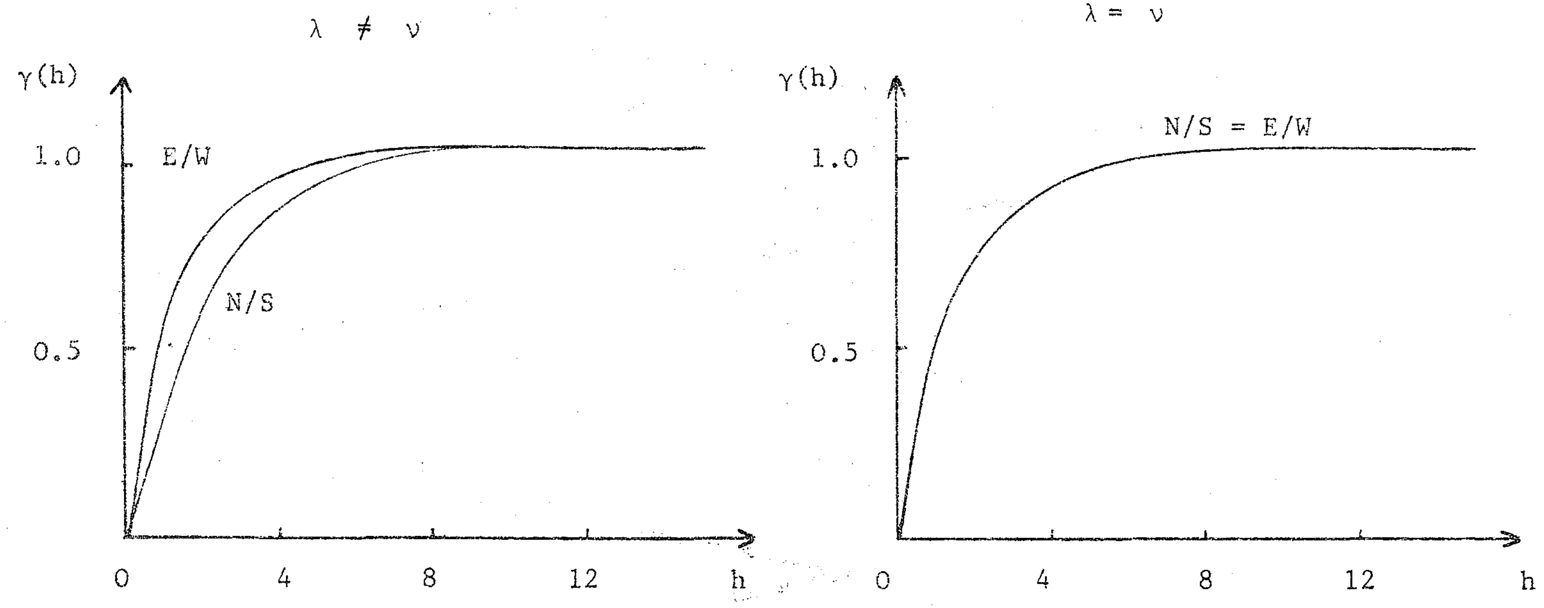


FIG.6.

TOP

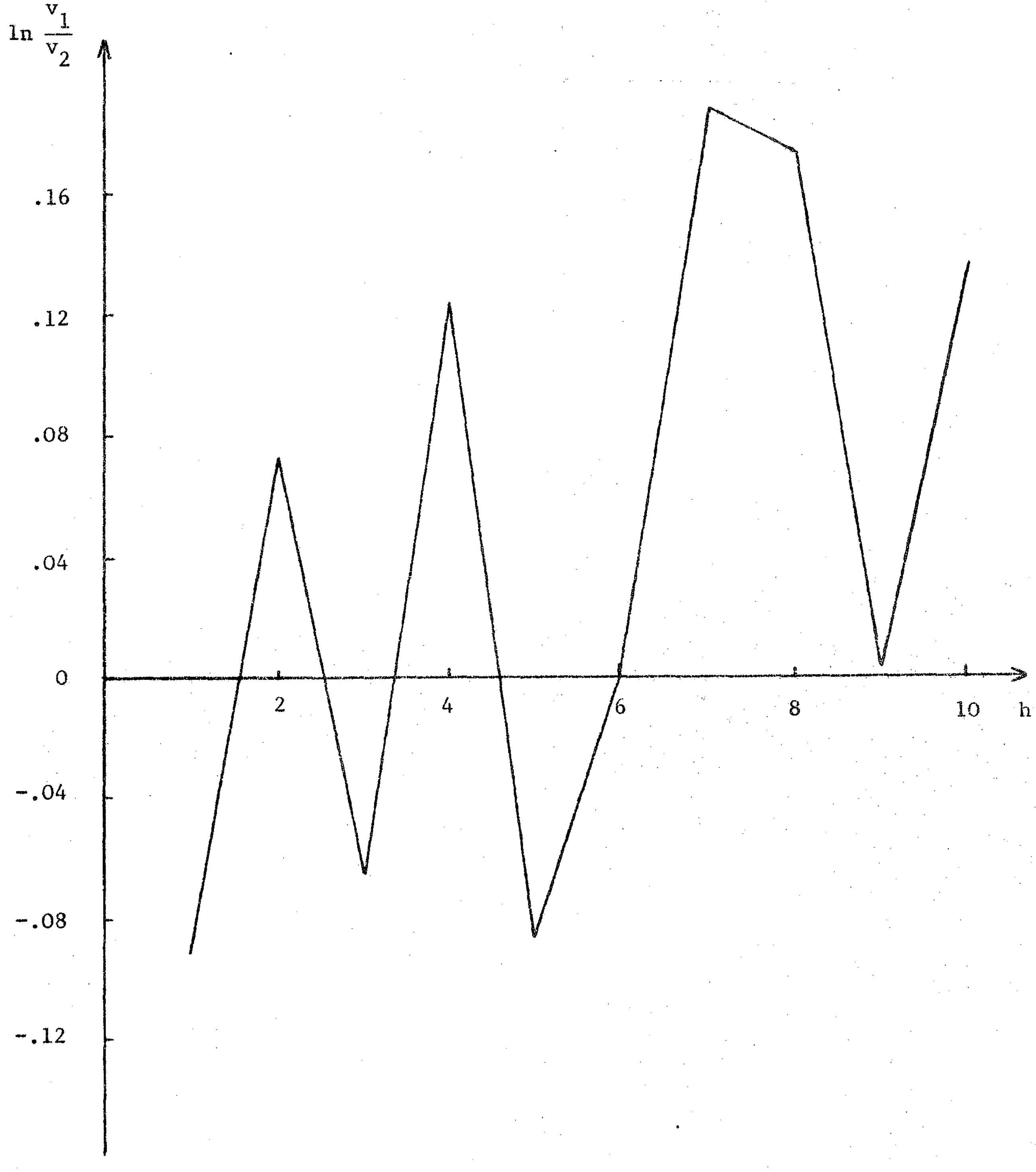


FIG.7.

TOP

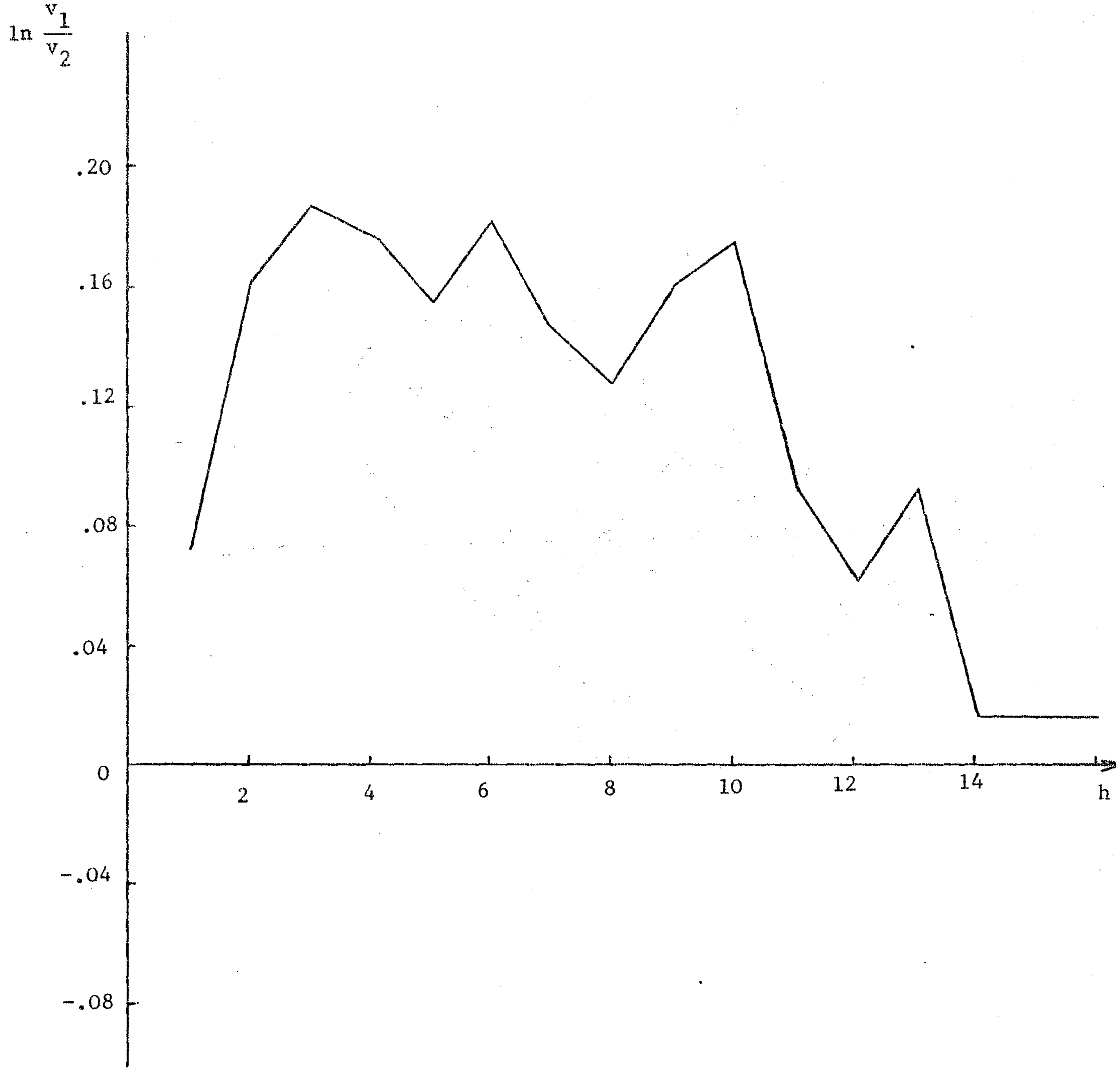


FIG.8.

TOP

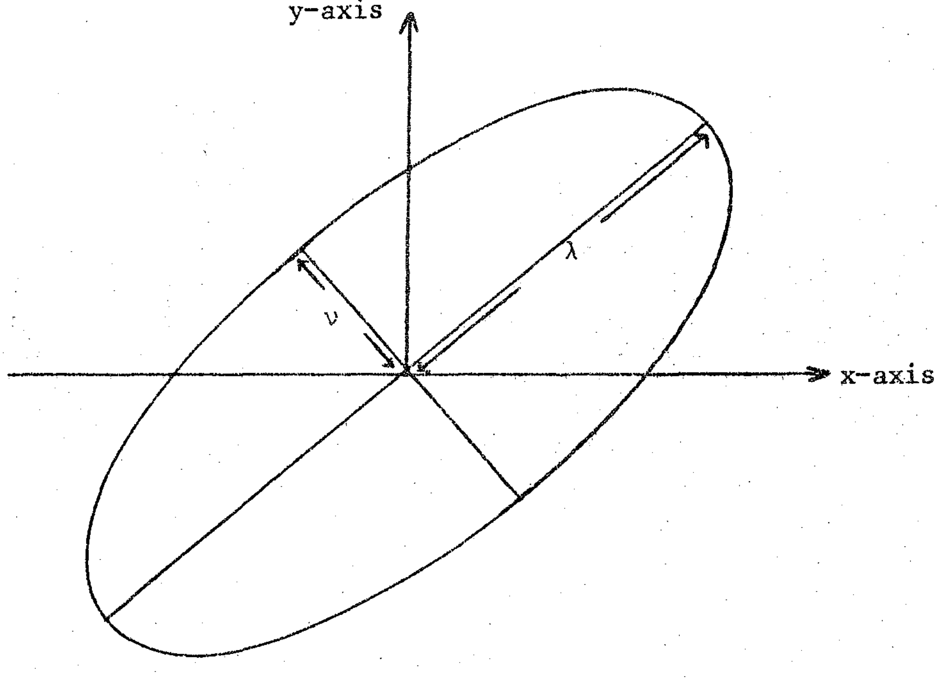


FIG.9.