

THE CALCULATION OF SEMI-VARIANCES  
IN THE TWO-DIMENSIONAL CASE

NOTE 63

J.V. Witter

Department of Hydraulics  
and Catchment Hydrology  
Agricultural University  
March 1983

## CONTENTS

	page
FOREWORD	1
1. INTRODUCTION	2
2. THE CALCULATION OF THE SAMPLE (SEMI-)VARIOGRAM	4
3. THE FITTING OF THE POPULATION (SEMI-)VARIOGRAM	6
4. THE USE OF THE PROGRAMS	8
4.1 The use of SAMVAR	8
4.2 The use of POPVAR	10
5. REFERENCES	11

## FIGURES

APPENDIX 1: LISTING OF SAMVAR.FOR

APPENDIX 2: LISTING OF POPVAR.FOR

## FOREWORD

This manual describes two computer programs: SAMVAR for the calculation of a sample semi-variogram and POPVAR for fitting a population semi-variogram. The programs are meant for use in the numerous cases (two-dimensional; no drift) in which application of the kriging method (Matheron; 1971) is simple. In complicated cases, the reader is referred to the computer programs listed in Journel and Huijbregts (1978) and in Kafritsas and Bras (1981).

Both programs include valuable suggestions made by M.A.J. van Montfort and by P.E.V. van Walsum. POPVAR also includes an IMSL subroutine. The programs have been in use for some years but are now generalized and documented.

## 1 INTRODUCTION

For a Regionalized variable  $Z(x)$ , which takes a value at every point  $x$  of coordinates  $(x_1, x_2)$  in two-dimensional space, the semi-variogram  $\gamma(h)$  is defined by:

$$\gamma(h) = \frac{1}{2} \text{Var} [Z(x) - Z(x+h)], \quad (1)$$

where  $h$  is distance.

$Z(x)$  should be a quantitative variable, measured on a linear scale.

If the so-called 'intrinsic hypothesis' holds:

$$\begin{cases} E [Z(x) - Z(x+h)] = 0, & (2a) \\ \text{Var} [Z(x) - Z(x+h)] = 2\gamma(h), & (2b) \end{cases}$$

then the population semi-variogram  $\gamma(h)$  is estimated without bias by the sample semi-variogram  $\hat{\gamma}(h)$ :

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i+h)]^2, \quad (3)$$

where  $z(x_i)$  are the experimental values ('realizations of  $Z(x)$ ') at points  $x_i$  such that data are available at  $x_i$  and  $x_i+h$ , and  $N(h)$  is the number of pairs of data points with a mutual distance  $h$ .

In case of a random sample the paired data are grouped according to distance classes, and  $N(h)$  is the number of pairs of data points in distance class  $h$ .

The semi-variogram is a tool:

- (a) to obtain insight in the structural properties of the Regionalized variable  $Z(x)$
- (b) to calculate estimation variances and to compare various interpolation schemes
- (c) in the application of kriging.

The influence of random errors on the estimation of the semi-variogram is given by Gandin (1965) and will not be discussed here. The following practical rules may be given concerning the estimation of the semivariogram (Journel and Huijbregts; 1978):  $N(h) \geq 30$  pairs, and the experimental semi-variogram should only be considered for small distances ( $h \leq L/2$ ) in relation to the dimension  $L$  of the domain on which it has been computed. The first rule follows from the well-known property (Kendall and Stuart; 1958. p.235) that in case of Normal distri-

buted increments  $[Z(x_1) - Z(x_1+h)]$  the variance of the estimator  $\underline{s}^2$  of the population variance  $\sigma^2$  equals:

$$\text{var } \underline{s}^2 \cong 2\sigma^4/N(h),$$

so that:

$$\frac{\text{var } \underline{s}^2}{(\sigma^2)^2} \cong \frac{2}{N(h)}. \quad (4)$$

So for a relative variance less than 5%,  $N(h)$  should exceed 40.

The second practical rule follows from the fact that very often only one realization of  $Z(x)$  is known. In case of a linear population semi-variogram

$$\gamma(h) = \alpha h,$$

where  $\alpha$  is a coefficient, and in the case this unique realization  $z(x)$  is known in all points of  $V$ , and with the help of all this information  $\gamma(h)$  was estimated by the estimator  $\hat{\gamma}'(h)$ , it can be shown (Matheron; 1971) that the relative variance of this estimator equals:

$$\frac{\text{var}(\hat{\gamma}'(h))}{[\gamma(h)]^2} = \begin{cases} \alpha^2 \left( \frac{4h^3}{3(L-h)} - \frac{h^4}{3(L-h)^2} \right), & (h \leq L/2) \quad (5a) \\ \alpha^2 \left( 2h^2 + \frac{1}{3}(L-h)^2 - \frac{4h}{3}(L-h) \right). & (h \geq L/2) \quad (5b) \end{cases}$$

So as soon as  $h$  is not very small compared with  $L$ , the relative variance becomes very large and statistical inference is no longer possible.

## 2 THE CALCULATION OF THE SAMPLE (SEMI-)VARIOGRAM

SAMVAR calculates values of the sample variogram according to (3). The distance scale and the distance classes have to be specified by the user. The specification of distance classes includes a guessing part: too large distance classes conceal the spatial structure, too narrow ones may give rise to pseudo fluctuations.

Sample variograms often exhibit strong fluctuations, see fig. 1 (from: Burrough et al.; 1983). These may be caused - apart from the chosen discretisation in distance classes - by anisotropy, by periodicities introduced because of the chosen sampling scale, because of 'real' periodicities or because of the skew distribution of the studied Regionalized variable. In case of real periodicities the frequencies in sample variograms of related variables should be equal. In case of a skewly distributed Regionalized variable the fluctuations may be reduced by sampling over a larger support (Journel and Huijbregts; 1978, p.77).

In hydrology often the 'multi-realization approach' is of interest: the studied Regionalized variable is sampled  $n$  times on the same set of data points. The  $n$  Individual Sample Variograms (ISV) are averaged to give one Multi-Realization Sample Variogram (MRSV). The reason for this practice will not be given here, reference is made to Chua and Bras (1980). It is interesting to note that there may be a very large spread of the  $n$  ISV's around the averaged MRSV, see fig. 2 (from: Witter; 1983). SAMVAR includes this possibility of averaging, and produces for the centre  $h$  of each distance class sample standard deviations  $s$  according to:

$$s = \left[ \left\{ \sum_{i=1}^n (\hat{\gamma}_i(h))^2 - n[\bar{\gamma}(h)]^2 \right\} / (n-1) \right]^{\frac{1}{2}}, \quad (6)$$

where:

$h$  = centre of a particular distance class

$\hat{\gamma}_i(h)$  = value of the  $i$ -th ISV at distance  $h$

$\bar{\gamma}(h)$  = value of the MRSV at distance  $h$ .

In the usual case of a single-realization approach also standard deviations  $s$  are calculated for each distance class, according to:

$$s = \left[ \left\{ \Sigma g^2 - [\hat{\gamma}(h)]^2 N(h) \right\} / (N(h)-1) \right]^{\frac{1}{2}}, \quad (7)$$

where:

$h$  = centre of a particular distance class

$N(h)$  = number of paired data in this distance class

$\hat{\gamma}(h)$  = value of the sample semi-variogram at distance  $h$

$\Sigma g^2 = \frac{1}{2} \Sigma [z(x_i) - z(x_j)]^2$  for all pairs  $x_i, x_j$  of data points  
in distance class with centre  $h$ .

The user may prefer to calculate semi-variances without pooling them into distance classes. In case of a single-realization approach no standard deviation can be calculated. In case of a multi-realization approach  $s$  is calculated according (6), but one should bear in mind that a 'distance class' in this case contains only the various repetitions over one pair of data points.

The standard deviations according to (6) or (7) give an impression of the spread of values around their averages.

### 3 THE FITTING OF THE POPULATION (SEMI-)VARIOGRAM

The fitting of a population variogram consists of the determination of certain parameters in a chosen variogram model (e.g.: linear). This may be done by eye and experience. Here a computer program is documented which fits variogram models to sample variograms by means of an error criterion (minimalization of the residual sums of squares).

Two models can be fitted by POPVAR. A linear model:

$$\gamma(h) = C\delta + \lambda_1 h, \quad (8)$$

where:

$h$  : distance

$C$  : a parameter for the 'nugget effect'

$\delta$  : 0 ( $h=0$ ) or 1 ( $h \neq 0$ )

$\lambda_1$ : parameter.

Also an exponential model can be fitted, according to:

$$\gamma(h) = C\delta + \lambda_1 (1 - \exp(-h/\lambda_2)), \quad (9)$$

where:

$\lambda_1$  = parameter (note: the 'sill' equals  $C+\lambda_1$ )

$\lambda_2$  = parameter (note: the 'range'  $\cong 3\lambda_2$ ).

If the semi-variogram is linear, the Regionalized variable  $Z(x)$  has an unlimited capacity for dispersion, and is only intrinsic. In case of an exponential semi-variogram the Regionalized variable  $Z(x)$  is second-order stationary.

The choice of these two models implies skipping of several frequently used alternatives:

(1) the logarithmic model:

$$\gamma(h) = C\delta + \alpha \ln h, \quad (10)$$

where:

$\alpha$  = parameter,

is discarded, because it cannot be used to describe regionalizations on a strict point support. Furthermore it is shown (Journel and Huijbregts; 1978) that the logarithmic model is the limit model of a nested succession of exponential models, the ranges of which increase geometrically.



(2) the Gaussian model:

$$\gamma(h) = C\delta + \lambda_1 [1 - \exp(-(h/\lambda_2)^2)], \quad (11)$$

is discarded, because properties exhibiting this very continuous and regular behaviour are seldom found in hydrological practice. In fact it is deterministic along any straight line in a plane: as shown by Yaglom (1962) for the corresponding correlation function the value of the process at an arbitrarily distant point can be approximated arbitrarily closely by a linear combination of past values of the process. A quasi Gaussian structure may be obtained in the case of errors in the coordinates (Journel and Huijbregts; 1978).

(3) the spherical model:

$$\gamma(h) = \begin{cases} C\delta + \left[ \frac{3}{2} \left(\frac{h}{\lambda_1}\right) - \frac{1}{2} \left(\frac{h}{\lambda_1}\right)^3 \right] \lambda_2, & (h \in [0, \lambda_1]) \\ C + \lambda_2, & (h \geq \lambda_1) \end{cases}$$

is discarded because it can very well be approximated by an exponential model, although this last model has a more smooth transition to the sill value (in fact it reaches its sill only asymptotically).

The user may discard, for reasons outlined in chapter one, all pairs of data at distances  $h > L/2$ . If a pair of data is used in the fitting procedure, no further weighting takes place. Fitting is preferably done by ordinary regression, because in the case of a very fluctuating sample variogram search methods give slightly different results in case of different starting values. In case of an exponential model a search method is used according to the Levenberg-Marquardt method (Abdy and Dempster; 1974), supplied by the IMSL algorithm ZXSSQ.

## 4 THE USE OF THE PROGRAMS

### 4.1 SAMVAR

Input to SAMVAR are two files:

- COR.DAT, containing the coordinates of the data points (FORMAT(2F)), and
- IN.DAT, containing on each line the values of the studied variables in the data point appearing in COR.DAT. on the same line (FORMAT (1X, 20 F6.1)).

The maximum number of data points is 100; the maximum number of variables is 20. Missing values are indicated by '999.0'.

The program is interactive. Questions are accompanied by the required formats for the answer on terminal. There are up to 15 distance classes of equal width. This width is determined by the parameter 'SCH', one of the parameters whose value is to be supplied on the terminal. If the largest distance between a pair of data points is L (in units of the coordinates), SCH may be chosen as 15/L. In case of SCH larger, e.g. SCH = 25/L, some distance classes are empty. In case of SCH smaller, the highest distance class (with the largest mean distance between points) contains relatively numerous pairs of data.

The following variables have to be supplied to SAMVAR on terminal:

NP : the number of data points

NV : the number of variables (or number of realizations if ISW = 1)

SCH: scaling factor

ISW: 1 (in case of a multi-realization approach) or 0 (otherwise)

IND: 1 (in case of optimization - with POPVAR - on semi-variances for each pair of data points) or 0 (in case of optimization on values of the semi-variogram for centres of distance classes).

So the program has four options:

- multi-realization approach vs. individual (ISW = 1, 0, respectively)
- fitting of the variogram model (with POPVAR) on semi-variances grouped to distance classes (IND = 0) or not (IND = 1).

Output on terminal is more or less regardless of the chosen option:

- the input on terminal is reproduced as a check (type Control C to interrupt the program, if necessary)

- for each variable coordinates and values for the first 5 data points are reproduced
- half the largest distance between data points is given. This value, HMAX, may be used in POPVAR.

The outputfile OUT.DAT contains:

- (1) if ISW = 0 and IND = 0 (test 1; app. I) the ISV's for all distance classes:  
 the first column contains a count for the variable, the second column a count for the distance class (empty distance classes are omitted), the third column the number of pairs of data points (should preferably be >40), the fourth column the mean distance, the fifth the value of ISV for this mean distance and the last column the standard deviation according to (7).
- (2) if ISW = 1 and IND = 0 (test 2; app. I) the single MRSV for all distance classes: the first column contains the number of the distance class (empty classes are omitted), the second column the number of pairs of data points for each ISV, underlying the MRSV, the third column the mean distance in a distance class, the fourth the values of MRSV and the last column the standard deviations according to (6).
- (3) if ISW = 0 and IND = 1 (test 3; app. I): semi-variances for each possible pair of data points. With NP datapoints there are ( in case of no missing values)  $NP(NP-1)/2$  such pairs. With  $NP < 100$ ,  $NP(NP-1)/2 < 4950$ . However, if  $NP(NP-1)/2$  large, the optimization with POPVAR demands excessive computer time, so it is recommended to let this number not exceed 100. In fact, if it does, both SAMVAR and POPVAR need some modifications.  
 The first column contains a count for the variable, the second column a count for the considered pair of data points, the third column gives their distance, the fourth column the semi-variance.
- (4) if ISW = 1 and IND = 1 (test 4; app. I): averaged semi-variances over the NV realizations for each possible pair of data points. The first column gives NV, the second gives the number of the paired data points used in the calculations, the third column their distance, the fourth column the semi-variance, and the fifth column the standard deviation according to (6).

4.2 POPVAR

Input to POPVAR is OUT.DAT, the output file of SAMVAR. The program is interactive. Questions are accompanied by the required formats for the answer on terminal.

The following variables have to be supplied to POPVAR on terminal:

- NO : the number of required optimizations (max:20)  
 IND : 1 (if optimization takes place on individual values) or 0 (otherwise)  
 HMAX: half of the largest distance between data points (i.e. HMAX from SAMVAR); if one wishes optimization on all data, HMAX should be given a large value  
 NC : number of distance classes ( $\leq 15$ ) or individual semi-variances ( $\leq 100$ )  
 ISW : 1 (if the linear semi-variogram model is chosen) or 0 (if the exponential model is preferred)  
 KX : 1 (if the nugget-effect is known) or 0 (if not)
- X(1),X(2) : starting values for  $\lambda_1, \lambda_2$  in (9)  
 X(1),X(2),X(3): starting values for C,  $\lambda_1, \lambda_2$  in (9)  
 X1 : the value of the nugget-variance (only if KX=1).

Output on terminal is self evident. If ISW = 0, the IMSL subroutine ZXSSQ is used. For the output of ZXSSQ the user should consult IMSL (1980).

Appendix 2 contains a listing of the program and a test run.

## 5 REFERENCES

- Abdy, P.R. and Dempster, M.A.H. (1974): 'Introduction to optimization methods' (Chapman and Hall; London).
- Burrough, P.A., Oerlemans, G., Stoffelsen, G. and Witter, J.V. (1983): 'A reconnaissance soil survey of the Hupselse Beek; Oost Gelderland' (Department of Hydraulics and Catchment Hydrology of the Agric. Univ., Wageningen).
- Chua, S.H. and Bras, R.L. (1980): 'Estimation of stationary and non-stationary random fields: kriging in the analysis of orographic precipitation' (MIT, Cambridge Mass.).
- Gandin, L.S. (1965): 'Objective analysis of meteorological fields' (Isr. Progr. for Scient. Transl., Jerusalem).
- IMSL (1980): 'The IMSL library Vol. 3' (Intern Math. & Stat. Libraries, Inc., Houston).
- Journel, A.G. and Huijbregts, Ch.J. (1978): 'Mining geostatistics' (Academic Press, London).
- Kafritsas, J. and Bras, R.L. (1981): 'The practice of kriging' (MIT, Cambridge Mass).
- Kendall, M.G. and Stuart, A (1958): 'The advanced theory of statistics, Vol. 1' (Charles Griffin, London).
- Matheron, G. (1971): 'The theory of regionalized variables and its applications' (École Nationale Supérieure des Mines de Paris, Paris).
- Witter, J.V. (1983): 'The areal reduction factor due to variability of rainfall' [in Dutch] (Department of Hydraulics and Catchment Hydrology of the Agric. Univ., Wageningen).
- Yaglom, A.M. (1962): 'An introduction to the theory of stationary random functions' (Prentice-Hall; Englewood Cliffs N.J.).

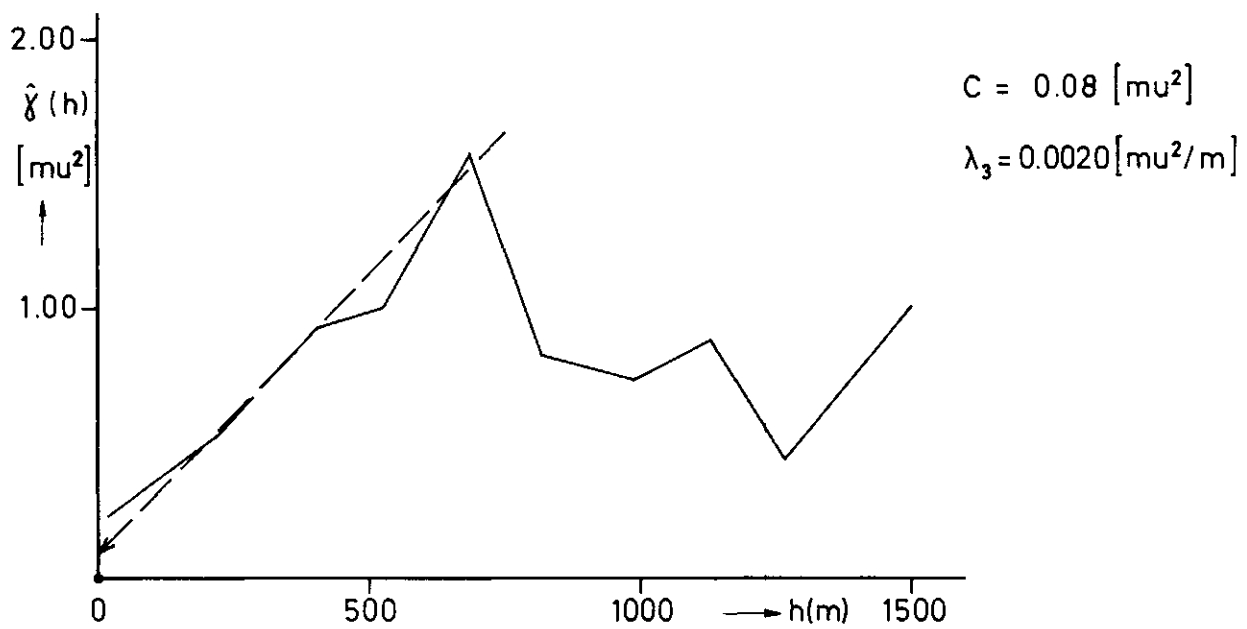
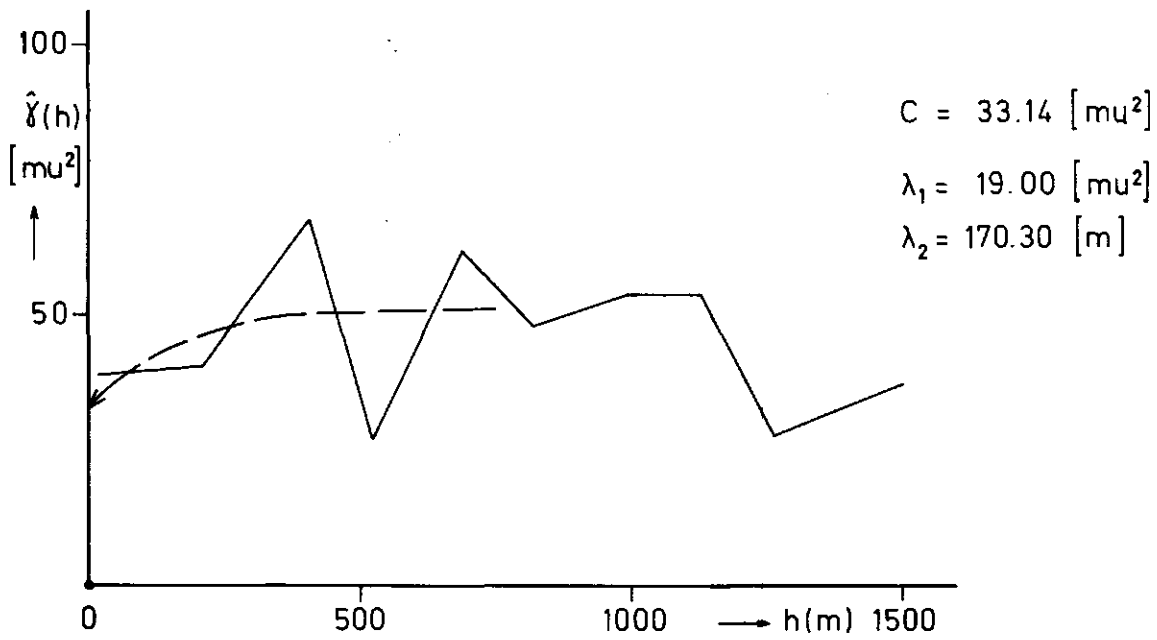


Fig. 1: Sample semi-variograms exhibiting fluctuations for some soil properties and fitted models with their parameters.

12

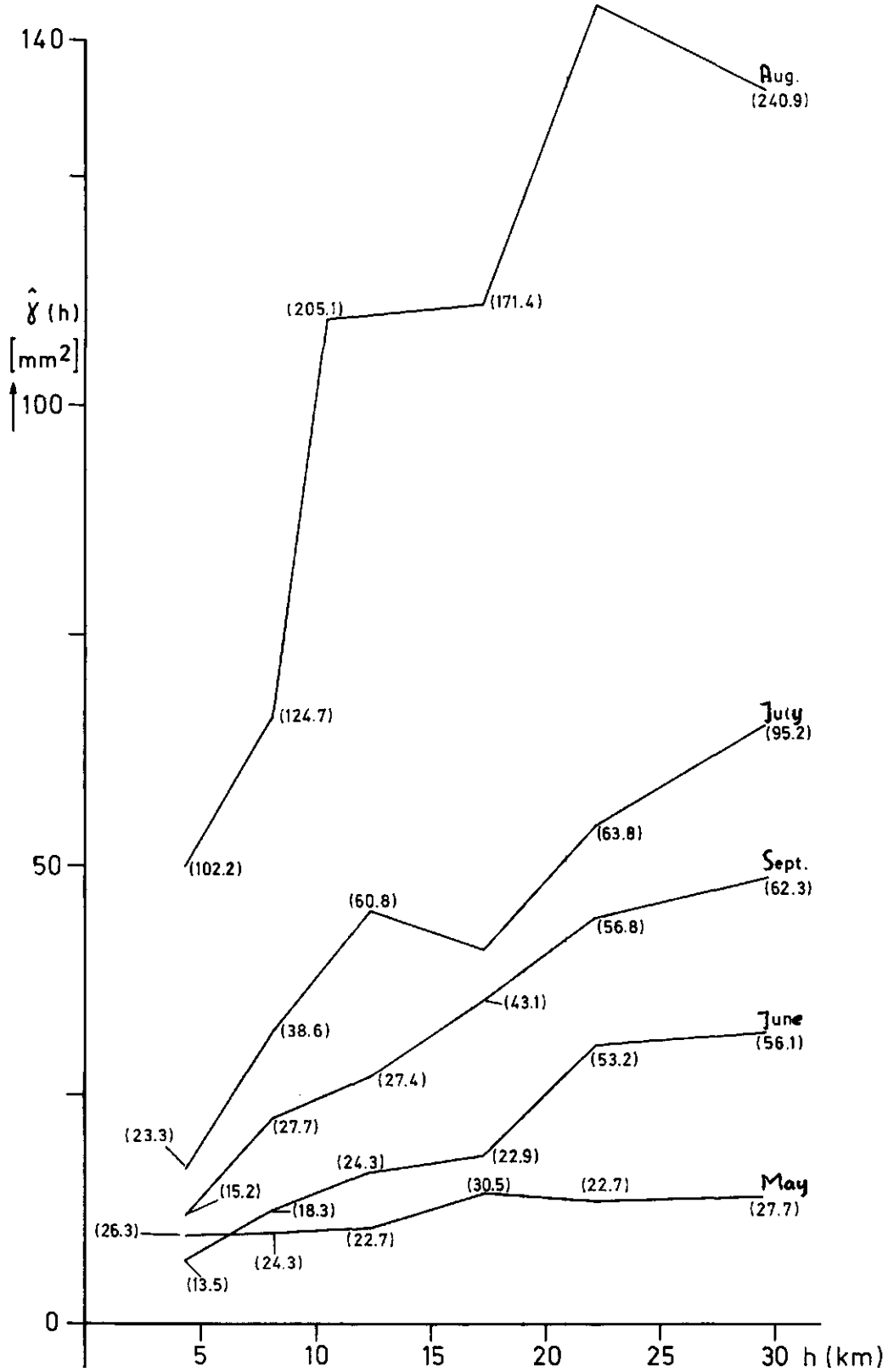


Fig. 2 : MRSV's for monthly maxima of daily summer rainfall. Numbers between brackets indicate the standard deviation of the ISV's around the MRSV.



13

Appendix I: LISTING OF SAMVAR.FOR

```

DIMENSION G1(15),G2(15),G3(15),G4(15),G5(15)
DIMENSION SD(15),SV(15),SUV(15)
INTEGER N(15)
DIMENSION SXV(100),SXV2(100),T(100),D3(100)
DIMENSION X(100),Y(100),P(100,20)
OPEN(UNIT=1,DEVICE='DSK',ACCESS='SEQIN',FILE='COR.DAT')
OPEN(UNIT=2,DEVICE='DSK',ACCESS='SEQIN',FILE='IN.DAT')
OPEN(UNIT=3,DEVICE='DSK',ACCESS='SEQOUT',FILE='OUT.DAT')
TYPE 100
FORMAT(/,' TYPE NP:NUMBER OF DATAPPOINTS(LE 100); FORMAT I3')
ACCEPT 101,NP
FORMAT(I3)
TYPE 102
FORMAT(/,' TYPE NV:NUMBER OF VARIABLES(LE 20); FORMAT I3')
ACCEPT 103,NV
FORMAT(I3)
TYPE 104
FORMAT(/,' TYPE SCH:SCALING FACTOR; FORMAT F10.2')
ACCEPT 105,SCH
FORMAT(F10.2)
DO 1 I=1,NP
READ(1,900) X(I),Y(I)
READ(2,901) (P(I,K),K=1,NV)
FORMAT(1X,2F)
900 FORMAT(1X,2F)
901 FORMAT(1X,20F6.1)
1 CONTINUE
TYPE 106
FORMAT(/,' MULTI-REAL.? YES:TYPE '1'; NO:TYPE '0'; FORMAT I1')
ACCEPT 107,ISM
FORMAT(I1)
TYPE 108
FORMAT(/,' OPTIM. ON IND.. VALUES (TYPE'1') OR NOT(TYPE'0')?')
ACCEPT 109,IND
FORMAT(I1)
TYPE 110
FORMAT(' AS A CHECK OUTPUT OF RESP. NP,NV,SCH,ISW,IND')
WRITE(5,905) NP,NV,SCH,ISW,IND
905 FORMAT(2I10,F10.2,2I10)
DO 16 I=1,100
T(I)=0.
16 C
C----- BEGIN MASTER LOOP
C
DO 2 ID=1,15
SXV(ID)=0.
SXV2(ID)=0.
T(ID)=0.
G1(ID)=0.
G2(ID)=0.
G3(ID)=0.
G4(ID)=0.
G5(ID)=0.
8
N(ID)=0.
SD(ID)=0.
SV(ID)=0.
SUV(ID)=0.
4 CONTINUE
GO TO 3
4
DO 9 I=1,NP
IF(I,LE,5) WRITE(5,902) I,X(I),Y(I),P(I,K)
DO 9 J=1,I
IF(J,EQ,I) GO TO 9
I3=I3+1
IF(P(I,K),EQ,999.) GO TO 9
IF(P(J,K),EQ,999.) GO TO 9
IF(I3,GT,100) WRITE(5,920)
920 FORMAT(' ERROR: NUMBER OF PAIRED DAT >100')
DD=SQRT((X(I)-X(J))**2+(Y(I)-Y(J))**2)
D3(I3)=DD
IF(DD,GT,DMAX) DMAX=DD
U=(P(I,K)-P(J,K))**2
XU=0.5#U
SXV(I3)=SXV(I3)+XU
SXV2(I3)=SXV2(I3)+XV*XV
T(I3)=T(I3)+1.
9 CONTINUE
D2=DMAX/2.
WRITE(5,907) K,D2
IF(ISW,EQ,1) GO TO 3
IE=I3
DO 10 I3=1,IE
WRITE(3,906) K,I3,D3(I3),SXV(I3)
906 FORMAT(2I6,6X,2F10.2)
SXV(I3)=0.
SXV2(I3)=0.
10 CONTINUE
3 IF(ISW,EQ,0) GO TO 11
IF(IND,EQ,1) GO TO 12
DO 13 ID=1,15
IF(G1(ID),EQ,0.) GO TO 13
XNV=G5(ID)
AN=G1(ID)/XNV
GSD=G2(ID)/XNV
GSV=G3(ID)/XNV
STD=SQRT((XNV*G4(ID)-G3(ID)*G3(ID))/(XNV*(XNV-1.)))
WRITE(3,921) ID,AN,GSD,GSV,STD
921 FORMAT(I6,F6.2,6X,3F10.2)
13 CONTINUE
GO TO 11
12 CONTINUE
I3=0
DO 14 I=1,NP
DO 14 J=1,I
IF(J,EQ,I) GO TO 14
I3=I3+1
XK2=T(I3)
IF(XK2,EQ,0.) GO TO 14
XK1=FLOAT(NV)
XK3=SQRT((X(I)-X(J))**2+(Y(I)-Y(J))**2)
XK4=SXV(I3)/XK2
STD=SQRT((XK2*SXV2(I3)-SXV(I3)*SXV(I3))/(XK2*(XK2-1.)))
WRITE(3,922) XK1,XK2,XK3,XK4,STD
922 FORMAT(F6.0,F6.2,6X,3F10.2)
14 CONTINUE
11 CONTINUE
END
TY COR.DAT
105.0 106.0
105.0 105.0
108.0 110.0
107.5 106.0
TY IN.DAT
25.0 26.0
27.0 27.0
34.0 29.0
20.0 28.0

```

16

TEST 1:

```

EXE SAMVAR.FOR
LINK: Loading
[LINKXCT SAMVAR execution]

TYPE NP:NUMBER OF DATAPPOINTS(LE 100); FORMAT I3
  4

TYPE NV:NUMBER OF VARIABLES(LE 20); FORMAT I3
  2

TYPE SCH:SCALING FACTOR; FORMAT F10.2
  2.50

MULTI-REAL.? YES:TYPE '1'; NO:TYPE '0'; FORMAT I1
  0

OPTIM. ON IND.. VALUES (TYPE'1') OR NOT(TYPE'0')?
  0
AS A CHECK OUTPUT OF RESP. NP,NV,SCH,ISW,IND
  4      2      2.50      0      0

```

```

*****
  1  105.00  106.00  25.00
  2  105.00  105.00  27.00
  3  108.00  110.00  34.00
  4  107.50  106.00  20.00
VARIABLE NO: 1 HMAX= 2.92

```

```

*****
  1  105.00  106.00  26.00
  2  105.00  105.00  27.00
  3  108.00  110.00  29.00
  4  107.50  106.00  28.00
VARIABLE NO: 2 HMAX= 2.92
CPU time 0.30 Elapsed time 45.62

```

```

EXIT

.TY OUT.DAT
  1  3  1  1.00  2.00  0.00
  1  7  2  2.60  18.50  8.9
  1  11 1  4.03  98.00  0.00
  1  13 1  5.00  40.50  0.00
  1  15 1  5.83  24.50  0.00
  2  3  1  1.00  0.50  0.00
  2  7  2  2.60  1.25  1.06
  2  11 1  4.03  0.50  0.00
  2  13 1  5.00  4.50  0.00
  2  15 1  5.83  2.00  0.00

```

TEST 2:

```

EXE SAMVAR.FOR
LINK: Loading
[LINKXCT SAMVAR execution]

TYPE NP:NUMBER OF DATAPPOINTS(LE 100); FORMAT I3
  4

TYPE NV:NUMBER OF VARIABLES(LE 20); FORMAT I3
  2

TYPE SCH:SCALING FACTOR; FORMAT F10.2
  2.50

MULTI-REAL.? YES:TYPE '1'; NO:TYPE '0'; FORMAT I1
  1

OPTIM. ON IND.. VALUES (TYPE'1') OR NOT(TYPE'0')?
  0
AS A CHECK OUTPUT OF RESP. NP,NV,SCH,ISW,IND
  4      2      2.50      1      0

```

```

*****
  1  105.00  106.00  25.00
  2  105.00  105.00  27.00
  3  108.00  110.00  34.00
  4  107.50  106.00  20.00
VARIABLE NO: 1 HMAX= 2.92

```

```

*****
  1  105.00  106.00  26.00
  2  105.00  105.00  27.00
  3  108.00  110.00  29.00
  4  107.50  106.00  28.00
VARIABLE NO: 2 HMAX= 2.92
CPU time 0.29 Elapsed time 44.38

```

```

EXIT

.TY OUT.DAT
  3  1.00  1.00  1.25  1.06
  7  2.00  2.60  9.88  12.20
  11 1.00  4.03  49.25  68.94
  13 1.00  5.00  22.50  25.46
  15 1.00  5.83  13.25  15.91

```

TEST 3:

```

EXE SAMVAR.FOR
LINK: Loading
[LINKXCT SAMVAR execution]

TYPE NP:NUMBER OF DATAPPOINTS(LE 100); FORMAT I3
  4

TYPE NV:NUMBER OF VARIABLES(LE 20); FORMAT I3
  2

TYPE SCH:SCALING FACTOR; FORMAT F10.2
  2.50

MULTI-REAL.? YES:TYPE '1'; NO:TYPE '0'; FORMAT I1
  0

OPTIM. ON IND.. VALUES (TYPE'1') OR NOT(TYPE'0')?
  1
AS A CHECK OUTPUT OF RESP. NP,NV,SCH,ISW,IND
  4      2      2.50      0      1

```

```

*****
  1  105.00  106.00  25.00
  2  105.00  105.00  27.00
  3  108.00  110.00  34.00
  4  107.50  106.00  20.00
VARIABLE NO: 1 HMAX= 2.92

```

```

*****
  1  105.00  106.00  26.00
  2  105.00  105.00  27.00
  3  108.00  110.00  29.00
  4  107.50  106.00  28.00
VARIABLE NO: 2 HMAX= 2.92
CPU time 0.30 Elapsed time 44.30

```

```

EXIT

.TY OUT.DAT
  1  1  1.00  2.00
  1  2  5.00  40.50
  1  3  5.83  24.50
  1  4  2.50  12.50
  1  5  2.69  24.50
  1  6  4.03  98.00
  2  1  1.00  0.50
  2  2  5.00  4.50
  2  3  5.83  2.00
  2  4  2.50  2.00
  2  5  2.69  0.50
  2  6  4.03  0.50

```

TEST 4:

```

EXE SAMVAR.FOR
LINK: Loading
[LINKXCT SAMVAR execution]

TYPE NP:NUMBER OF DATAPPOINTS(LE 100); FORMAT I3
  4

TYPE NV:NUMBER OF VARIABLES(LE 20); FORMAT I3
  2

TYPE SCH:SCALING FACTOR; FORMAT F10.2
  2.50

MULTI-REAL.? YES:TYPE '1'; NO:TYPE '0'; FORMAT I1
  1

OPTIM. ON IND.. VALUES (TYPE'1') OR NOT(TYPE'0')?
  1
AS A CHECK OUTPUT OF RESP. NP,NV,SCH,ISW,IND
  4      2      2.50      1      1

```

```

*****
  1  105.00  106.00  25.00
  2  105.00  105.00  27.00
  3  108.00  110.00  34.00
  4  107.50  106.00  20.00
VARIABLE NO: 1 HMAX= 2.92

```

```

*****
  1  105.00  106.00  26.00
  2  105.00  105.00  27.00
  3  108.00  110.00  29.00
  4  107.50  106.00  28.00
VARIABLE NO: 2 HMAX= 2.92
CPU time 0.28 Elapsed time 50.44

```

```

EXIT

.TY OUT.DAT
  2  2.00  1.00  1.25  1.06
  2  2.00  5.00  22.50  25.46
  2  2.00  5.83  13.25  15.91
  2  2.00  2.50  7.25  7.42
  2  2.00  2.69  12.50  16.97
  2  2.00  4.03  49.25  68.94

```

15

Appendix II: LISTING OF POPVAR.FOR

```

EXTERNAL EXMPL2,EXMPL3
INTEGER M,N,IXJAC,NSIG,MAXFN,IOPT,INFER,IER
REAL PARM(4),X(3),F(100),XJAC(100,3),XJTJ(6),WORK(221),EPS,
1 DELTA,SGN,SSQ,Y(100),D(100)
COMMON /ZSQ/ Y,D
COMMON /XX3/X1
OPEN(UNIT=1,DEVICE='DSK',ACCESS='SEQIN',FILE='OUT.DAT')
TYPE 100
100 FORMAT(/,' TYPE NO:NUMBER OF OPTIMIZATIONS (LE 20);FORMAT I2')
ACCEPT 101,NO
101 FORMAT(I2)
TYPE 102
102 FORMAT(/,' OPT. ON IND. VALUES<TYPE'1'> OR NOT<TYPE'0'>?')
ACCEPT 103,IND
103 FORMAT(I1)
TYPE 104
104 FORMAT(/,' TYPE VALUE OF .5*DMAX; FORMAT F10.2')
ACCEPT 105,XL2
105 FORMAT(F10.2)
C
C----- START MASTER LOOP
C
DO 1 N1=1,NO
TYPE 120
120 FORMAT(/,' *****')
NSIG=5
EPS=0.000001
DELTA=0.0
MAXFN=100000
IOPT=1
TYPE 106
106 FORMAT(/,' TYPE NC: NUMBER OF DIST. CLASSES ; FORMAT I2')
ACCEPT 107,NC
107 FORMAT(I2)
READ(1,900) ((D(I),Y(I)),I=1,NC)
TYPE 118
118 FORMAT(/,' AS A CHECK, OUTPUT OF FIRST VALUE OF D(.) AND Y(.)')
WRITE(5,900) D(1),Y(1)
900 FORMAT(18X,2F10.2)
I2=0
DO 6 I=1,NC
IF(D(I).GT.XL2) GO TO 6
I2=I2+1
D(I2)=D(I)
Y(I2)=Y(I)
6 CONTINUE
NC=I2
TYPE 108
108 FORMAT(/,' LIN. <TYPE'1'> OR EXP. <TYPE'0'> VARIOGRAM?')
ACCEPT 109,ISM
109 FORMAT(I1)
TYPE 110
110 FORMAT(/,' NUGGET EFFECT KNOWN <TYPE'1'> OR NOT <TYPE'0'>?')
ACCEPT 111,KX
111 FORMAT(I1)
IF(19W.EQ.1) GO TO 50
IF(KX.EQ.0) GO TO 10
C
C----- EXPONENTIAL VARIOGRAM; NUGGET-EFFECT KNOWN
C
TYPE 112
112 FORMAT(/,' TYPE STARTING VALUES FOR X(1) AND X(2);FORMAT 2F')
ACCEPT 113,X(1),X(2)
113 FORMAT(2F)
M=NC
N=2
IXJAC=100
TYPE 114
114 FORMAT(/,' TYPE X1;NUGGET EFFECT; FORMAT F')
ACCEPT 115,X1
115 FORMAT(F10.2)
WRITE(5,901) ((D(K),Y(K)),K=1,NC)
901 FORMAT(1H,'AFSTAND=',F6.2,'SEMI-VAR.',F10.3)
CALL ZXSSQ(EXMPL2,M,N,NSIG,EPS,DELTA,MAXFN,IOPT,PARM,X,SSQ,F,
1 XJAC,IXJAC,XJTJ,WORK,INFER,IER)
WRITE(5,902) X(1),X(2)
902 FORMAT(1H,'ESTIMATES OF PARAMETERS:','1H,'X(1)='F14.6,
1 'X(2)='F10.6/')
WRITE(5,903) SSQ,IER,INFER
903 FORMAT(1H,'SSQ='F14.3,' IER='I5,' INFER='I5,/)
WRITE(5,904) (WORK(I),I=1,5)
904 FORMAT(1H,'NORM OF GRADIENT='F12.3,' FUNCTION EV.'F7.0,
1 1H,'SIGN.DIGITS='F5.0,' MARQUARDT PARAMETER='F10.5,/,1H,
1 'ITERATIONS='F5.0,/)
GO TO 1
C
C----- EXPONENTIAL VARIOGRAM; NUGGET-EFFECT NOT KNOWN
C
TYPE 116
116 FORMAT(/,' TYPE STARTINGVALUES FOR X(1),X(2),X(3); FORMAT 3F')
ACCEPT 117,X(1),X(2),X(3)
117 FORMAT(3F)
M=NC
N=3
IXJAC=100
WRITE(5,901) ((D(K),Y(K)),K=1,NC)
CALL ZXSSQ(EXMPL3,M,N,NSIG,EPS,DELTA,MAXFN,IOPT,PARM,X,SSQ,F,
1 XJAC,IXJAC,XJTJ,WORK,INFER,IER)
WRITE(5,905) X(1),X(2),X(3)
905 FORMAT(' ESTIMATES OF PARAMETERS:','/
1 'X(1)='F14.6,' X(2)='F14.6,' X(3)='F14.6/')
WRITE(5,903) SSQ,IER,INFER
WRITE(5,904) (WORK(I),I=1,5)
GO TO 1
C
C----- LINEAR VARIOGRAM; NUGGET-EFFECT KNOWN
C
50 CONTINUE
IF(KX.EQ.0) GO TO 60
TYPE 114
ACCEPT 115,X1
SDY=0.
SD2=0.
DO 2 I=1,NC
YI=Y(I)-X1
SDY=SDY+D(I)*YI
SD2=SD2+D(I)*D(I)
B=SDY/SD2
WRITE(5,906) B
906 FORMAT(' ESTIMATE OF PARAMETER:','/
1 'B='F14.6/')
GO TO 1
C
C----- LINEAR VARIOGRAM; NUGGET-EFFECT NOT KNOWN
C
60 CONTINUE
SDY=0.
SD=0.
SD2=0.
SY=0.
XN=FLOAT(NC)
DO 3 I=1,NC
SDY=SDY+D(I)*Y(I)
SY=SY+Y(I)
SD=SD+D(I)
SD2=SD2+D(I)*D(I)
B=(SDY-SD*SY/XN)/(SD2-SD*SD/XN)
A=(SY-B*SD)/XN
WRITE(5,907) A,B
907 FORMAT(' ESTIMATES OF PARAMETERS:','/
1 'A='F14.6,' B='F14.6/')
1 CONTINUE
END
SUBROUTINE EXMPL2(X,M,N,F)
INTEGER M,N,I
REAL X(N),F(M),Y(100),D(100)
COMMON /ZSQ/ Y,D
COMMON /XX3/X1
SS=0.
DO 1 I=1,M
F(I)=Y(I)-X(1)-X(2)*(1.-EXP(-D(I)/X(2)))
SS=SS+F(I)**2.
RETURN
END
SUBROUTINE EXMPL3(X,M,N,F)
INTEGER M,N,I
REAL X(N),F(M),Y(100),D(100)
COMMON /ZSQ/ Y,D
COMMON /XX3/X1
SS=0.
DO 1 I=1,M
F(I)=Y(I)-X(1)-X(2)*(1.-EXP(-D(I)/X(3)))
SS=SS+F(I)**2.
RETURN
END

```

.TY OUT.DAT

1	64	20.44	22.25
4	46	180.48	65.18
5	121	224.61	117.32
6	25	262.64	92.00
7	15	336.90	71.03
8	25	374.12	41.02
9	65	432.70	35.64
10	51	472.03	67.15
11	44	523.35	40.39
12	52	574.98	43.65
1	64	20.44	5.61
4	46	180.48	252.90
5	121	224.61	177.89
6	25	262.64	582.00
7	15	336.90	503.30
8	25	374.12	107.94
9	65	432.70	320.92
10	51	472.03	185.44
11	44	523.35	339.52
12	52	574.98	114.42

16

TEST:

```

EXE POPVAR.FOR,IMSL.REL/LIB
FORTRAN: POPVAR
MAIN.
EXMPL2
EXMPL3
LINK: Loadins
[LNKXCT POPVAR execution]

TYPE NO:NUMBER OF OPTIMIZATIONS (LE 20);FORMAT I2
 2

OPT. ON IND. VALUES(TYPE*1*) OR NOT(TYPE*0*)?
0

TYPE VALUE OF .5*DMAX; FORMAT F10.2
 375.00

*****
TYPE NC: NUMBER OF DIST. CLASSES ; FORMAT I2
10

AS A CHECK, OUTPUT OF FIRST VALUE OF D(.) AND Y(.)
 20.44  22.25

LIN. (TYPE*1*) OR EXP. (TYPE*0*) VARIOGRAM?
0

NUGGET EFFECT KNOWN (TYPE*1*) OR NOT (TYPE*0*)?
1

TYPE STARTING VALUES FOR X(1) AND X(2);FORMAT 2F
50. 57.

TYPE X1:NUGGET EFFECT; FORMAT F
17.66
AFSTAND= 20.44SEMI-VAR.= 22.250
AFSTAND=180.48SEMI-VAR.= 65.180
AFSTAND=224.61SEMI-VAR.= 117.320
AFSTAND=262.64SEMI-VAR.= 92.000
AFSTAND=336.90SEMI-VAR.= 71.030
AFSTAND=374.12SEMI-VAR.= 41.020
ESTIMATES OF PARAMETERS:
X(1)= 60.107918 X(2)= 63.039370

SSQ= 3541.563 IER= 0 INFER= 2

NORM OF GRADIENT= 0.307 FUNCTION EV.= 31.
SIGN.DIGITS= 4. MARQUARDT PARAMETER= 2.66087
ITERATIONS= 6.

*****
TYPE NC: NUMBER OF DIST. CLASSES ; FORMAT I2

```

```

10
AS A CHECK, OUTPUT OF FIRST VALUE OF D(.) AND Y(.)
 20.44  5.61

LIN. (TYPE*1*) OR EXP. (TYPE*0*) VARIOGRAM?
0

NUGGET EFFECT KNOWN (TYPE*1*) OR NOT (TYPE*0*)?
1

TYPE STARTING VALUES FOR X(1) AND X(2);FORMAT 2F
250. 100.

TYPE X1:NUGGET EFFECT; FORMAT F
2.09
AFSTAND= 20.44SEMI-VAR.= 5.610
AFSTAND=180.48SEMI-VAR.= 252.900
AFSTAND=224.61SEMI-VAR.= 177.890
AFSTAND=262.64SEMI-VAR.= 582.000
AFSTAND=336.90SEMI-VAR.= 503.300
AFSTAND=374.12SEMI-VAR.= 107.940
ESTIMATES OF PARAMETERS:
X(1)= 365.739956 X(2)=121.595228

SSQ= 171197.389 IER= 0 INFER= 2

NORM OF GRADIENT= 1.965 FUNCTION EV.= 24.
SIGN.DIGITS= 2. MARQUARDT PARAMETER= 0.03872
ITERATIONS= 9.

CPU time 0.43 Elapsed time 2:25.64

EXIT

```



17