

ON FITTING THE MICHAELIS-MENTEN EQUATION
FOR ENZYME KINETICS USING
GENERALIZED LINEAR MODELS*

BU-646-M

by

May, 1978

H. V. Henderson¹ and H. E. Walls²

Biometrics Unit, Cornell University, Ithaca, New York

Abstract

The Michaelis-Menten equation, the rate equation for a one-substrate enzyme-catalyzed reaction, is an inverse polynomial, and may be fitted using a generalized linear model with inverse link and normal error. A GENSTAT program is given that performs this analysis for an arbitrary number of sets each with a varying number of subsets which have a varying number of observations. Output for one set is given, which includes a graph of the data points on the measurement and reciprocal scales (on which the relationship is linear), graphs of residuals and kinetic parameter estimates for the model for the set and each of its subsets, with a test of homogeneity of the subsets.

1. Introduction

The Michaelis-Menten equation has proved to be a useful model in enzyme kinetics, see for example Cleland [1970], and was developed by Michaelis and Menten [1913] as the rate equation for a one-substrate enzyme-catalyzed reaction. It relates the initial velocity of reaction, V_0 , the maximum velocity, V_{\max} , and the initial substrate concentration, S , through the Michaelis-Menten constant, K_m , by

$$V_0 = \frac{V_{\max} S}{K_m + S} \quad (1)$$

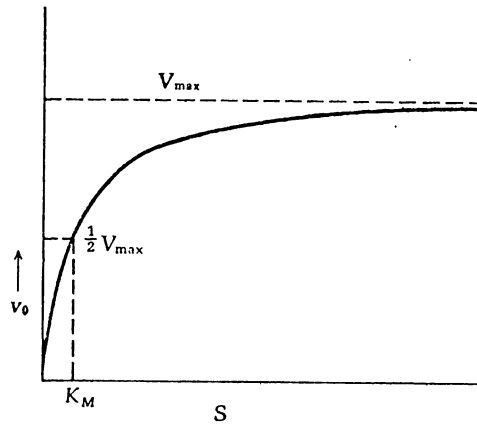
* Paper No. BU-646-M in the Biometrics Unit Mimeo Series, Cornell University.

1 Supported by a New Zealand National Research Advisory Council Fellowship with computing support from New York State College of Agriculture, Cornell University.

2 Department Physiology, Biochemistry and Pharmacology, N.Y.S. College of Veterinary Medicine, Cornell University; supported by NIH grant AM-19625.

K_m , which is the substrate concentration corresponding to an initial velocity of $\frac{1}{2} v_{max}$, and v_{max} are the parameters of the model which characterise the reaction.

Figure 1: Michaelis-Menten equation



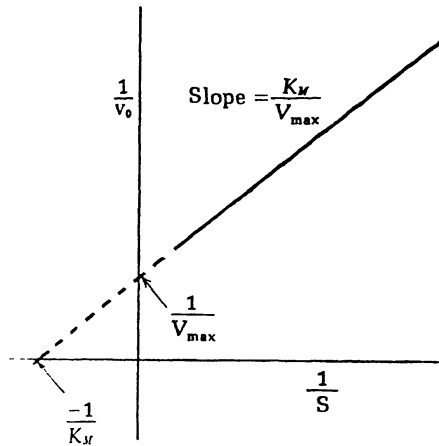
$$v_0 = \frac{v_{max} S}{K_m + S}$$

The Michaelis-Menten equation is nonlinear on the measurement scale. Methods of estimating the parameters have been by graphical techniques applied to transformations of the data. One common transformation is derived by taking the reciprocal of both sides of (1) giving the Lineweaver-Burk equation,

$$\frac{1}{v_0} = \frac{K_{max}}{v_{max}} \cdot \frac{1}{S} + \frac{1}{v_{max}}, \quad (2)$$

which is linear on the reciprocal scale. Lineweaver and Burk [1934] introduced the double-reciprocal plot (figure 2), of plotting $1/v_0$ against $1/S$, to estimate the straight line that 'best' fits the data points. Extrapolating this line yields a $\frac{1}{v_0}$ - intercept, which estimates $1/v_{max}$, and a $\frac{1}{S}$ - intercept, which estimates $-1/K_m$.

Figure 2: Lineweaver-Burk equation



A recent refinement of the double-plot technique, for estimating K_m and V_{max} , is by using the method of least squares to fit the straight line regression, as the 'best' straight line,

$$\frac{1}{V_0} = \text{slope} \cdot \frac{1}{S} + \text{intercept} + \text{error}, \quad (3)$$

with parameter estimates, denoted by a $\hat{}$ over the parameter,

$$\begin{aligned} \hat{V}_{max} &= 1 / \text{intercept} \\ \hat{K}_m &= \text{slope} / \text{intercept}. \end{aligned} \quad (4)$$

Assumptions implicit in this method include:

- 1A) The error structure is additive on the reciprocal scale.
- B) The errors are Independently Identically Distributed (IID) with mean 0 and constant variance over observations within the 1 set of experiments.
- C) The error is normally distributed for testing and Maximum Likelihood Estimation (MLE) of the slope and intercept so that \hat{K}_m and \hat{V}_{max} are also MLE.
- D) S and hence 1/S are measured without error.

2. Generalized Linear Model Approach

A more realistic assumption on the error structure, instead of 1A), is :

2A) The error is additive on the measurement, V , scale.

Then with this assumption the model can be rewritten as

$$V_o = \text{reciprocal} \left(\text{slope} \cdot \frac{1}{S} + \text{intercept} \right) + \text{error}. \quad (5)$$

This is a special case of generalized linear models, discussed by Nelder [1968] and Nelder and Wedderburn [1972]. This class of models explains the observations, \underline{y} , as a combination of a systematic component, \underline{m} , and an error component, \underline{e} , following a specified distribution, so that $\underline{y} = \underline{m} + \underline{e}$. Linear model theory with $\underline{m} = \underline{Xb}$, where \underline{X} is a design matrix and \underline{b} is the parameter vector, is generalized to $\underline{m} = f(\underline{Xb})$, a function of the linear predictor. The specification of the model is made in terms of the "link function" which is the inverse of f , so that $\underline{Xb} = \text{link}(\underline{m})$. The algorithm gives the maximum likelihood estimate of \underline{b} , under the assumed probability distribution, by iterative weighted least squares. We fit (1) in the form (5) with an inverse link and normal error, obtaining the MLE of V_{\max} and K_m by (4).

3. GENSTAT Program

A GENSTAT program was written to fit (5) for an arbitrary number of sets each with a varying number of subsets, which have a varying number of observations. The analysis for each set is achieved by first regressing V_o , with an inverse link, on $1/S$ so that a common slope and intercept for each subset is specified. The model is then modified, so that each subset has a separate slope and intercept, to give the parameter estimates for each subset.

GENSTAT is a very flexible statistical language and we exploit this by setting up a MACRO, or subroutine, to perform the analysis for each set, with the sizes of the structures changing dynamically for each set. Results from one procedure may be communicated to another so that data and residual plots are easily obtained.

The actual data set analysed had 84 sets with from 1 to 7 subsets and was run without residual plots in 4 runs using GENSTAT 4.01, each producing around 5,000 lines, requiring a maximum region of 184K and taking 25 seconds CPU time on the Cornell IBM 370/168.

The enzyme studied was microsomal D-glucose-6-phosphate phosphohydrolase (E.C.3.1.3.9) obtained from rat liver. The rate of glucose-6-phosphate hydrolysis was determined (in μ moles inorganic phosphate formed /min/mg microsomal protein) at 4 substrate concentrations in the presence of a variety of buffers at 2 pH values. The rationale for the kinetic analysis was that several of the buffers were suspected to be inhibitors of the enzyme, and their effects should be reflected by changes in the kinetic parameters, V_{max} and/or K_m . (Some of this work appeared in Walls [1978])

The program listing and sample output from one set with 7 subsets, annotated where appropriate, follows. The subset number is used as the plotting symbol on the graphs.

•REFERENCE/LINE=5000• ENZYME_KINETICS
•PAGE•
•CAPTION•
..

THE INITIAL VELOCITY OF REACTION, V,
IS MEASURED FOR GIVEN SUBSTRATE CONCENTRATION, S.
THE MICHAELIS-MENTEN PREDICTION EQUATION

$$V = V_{MAX} * S / (K_M + S)$$

IS THE ASSUMED STRUCTURAL BASIS FOR ESTIMATING
V_{MAX} AND K_M.
WHICH CHARACTERIZE THE REACTION.

TO DO THIS AN ERROR STRUCTURE MUST BE ASSUMED.

1) THE USUAL METHOD IS TO TRANSFORM THE
MICHAELIS AND MENTEN EQUATION, BY TAKING RECIPRICALS TO GIVE

$$RV = (K_M / V_{MAX}) * RS + 1 / V_{MAX}$$

WHICH IS LINEAR ON THE RECIPROCAL SCALE.

FURTHER ASSUMING

- 1A) THE ERROR STRUCTURE IS ADDITIVE
ON THE RECIPROCAL SCALE
- B) THE ERRORS ARE INDEPENDENTLY IDENTICALLY DISTIBUTED
(IID) WITH MEAN 0 AND CONSTANT VARIANCE
OVER OBSERVATIONS WITHIN THE 1 SET OF EXPERIMENTS
- C) THE ERROR IS NORMALLY DISTRIBUTED
FOR TESTING AND MAXIMUM LIKELIHOOD ESTIMATION (MLE)
OF THE SLOPE AND INTERCEPT SO THAT K_M AND V_{MAX}
ARE ALSO MLE
- D) S AND HENCE 1/S ARE MEASURED WITHOUT ERROR.

THEN A STRAIGHT LINE REGRESSION OF

$$RV = SLOPE * RS + INTERCEPT + ERROR$$

GIVES ESTIMATES

$$V_{MAX} = 1 / INTERCEPT$$
$$K_M = SLOPE / INTERCEPT.$$

2) A MORE REALISTIC ASSUMPTION ON THE
ERROR STRUCTURE, INSTEAD OF 1A), IS

2A) THE ERROR IS ADDITIVE ON THE V SCALE.

WITH THIS ASSUMPTION THE MODEL

$$V = RECIPROCAL (SLOPE * RS + INTERCEPT) + ERROR$$

IS FITTED USING THE GENERAL FRAMEWORK OF THE
GENERALISED LINEAR MODEL WITH NORMAL ERROR
AND RECIPROCAL LINK.
..

•PAGE•
•SCALAR• NSETS
•READ• NSETS
•CAPTION•

••
THE ANALYSIS IS DONE FOR EACH OF THE NSETS SETS
••

•SCALAR• I = 1

•MACRO• ANALYSIS \$

•PAGE•

••
THE MACRO "ANALYSIS" PERFORMS THE ESTIMATION OF VMAX AND KM FOR EACH SET AND ITS SUBSETS, ACCEPTING A VARYING NUMBER OF SETS AND OBSERVATIONS WITHIN EACH SUBSET. THE MICHAELIS-MENTEN PREDICTION EQUATION IS FITTED USING AN INVERSE LINK AND NORMAL ERROR IN THE GENERAL FRAMEWORK OF THE GENERALISED LINEAR MODEL.
••

•START•

•SCALAR• NSUBSETS, SLOPE, INTERCEPT, KM, VMAXI, SET, SUBSET
•READ/P• SET, NSUBSETS
••

SUBSET S AND V VALUES ARE READ IN TURN FOR EACH SET AND THEIR VARIABLE LENGTH IS DETERMINED BY THE NUMBER OF VALUES READ
••

•READ/P,NUN=V• SUB_SET, S, V \$ F, 1, 3.1, 4.3, /

•SCALAR• LENGTH, NCOEFS ,J

•RUN•

•START•

•CALC• LENGTH = NVAL(S)

•CALC• NCOEFS = 2 * NSUBSETS

•RUN•

•START•

•VARIATE• RV, RS , VHAT, RESV, VHATSEP, RESVSEP \$ LENGTH

•VARIATE• SEPCOEFS \$ NCOEFS

•FACTOR• SUBSETS \$ NSUBSETS, LENGTH

•RUN•

•CALC• RV = 1 / V

: RS = 1 / S

•CAPTION•

••
THE DATA SET PLUS RS = 1/S AND RV = 1/V FOR
••

•PRINT/P• SET \$ 8

•PRINT/P• SUB_SET, S, V, RS, RV \$ 18, 18.1, 3(18.3)

•GROUPS• SUBSETS = INTPT (SUB_SET)

•HEADING• V_AXIS = ••V INITIAL VELOCITY OF REACTION••

: S_AXIS = ••S SUBSTRATE CONCENTRATION••

•GRAPH/ATY=V_AXIS, ATX=S_AXIS,NRF=100• V ; S \$; SUBSETS

•PRINT/P• SET \$ 8

•HEADING• RV_AXIS = ••RV RECIPROCAL INITIAL VELOCITY REACTION••

: RS_AXIS = ••RS RECIPROCAL SUBSTRATE CONCENTRATION••

•GRAPH/ ATY=RV_AXIS, ATX=RS_AXIS,NRF=100• RV ; RS \$; SUBSETS

•PRINT/P• SET \$ 8

•PAGE•


```

*CALC*           J = 1
*LABEL*         NEXT_SUBSET
  *CALC*         INTERCEPT = ELEM ( SEPCOEFS ; J )
  *CALC*         SLOPE       = ELEM ( SEPCOEFS ; NSUBSETS + J )
  :             VMAXI        = 1 / INTERCEPT
  :             KM           = SLOPE / INTERCEPT
  :             SUBSET       = J
*PRINT/P*       SUBSET, VMAXI, KM   $ 10, 2(18.9)
*CALC*         J = J + 1
*JUMP*         NEXT_SUBSET * ( J .LE. NSUBSETS )
*CAPTION*

```

```

-----
*HEADING*      V_HATSEP = **VHATSEP  PRED SEPARATE LINES FOR SUBSETS**
:              RES_VSEP = **RESVSEP  FROM LINES FOR EACH SUBSET**
*GRAPH/ATY=RES_VSEP,ATX=V_HATSEP,NRF=100*RESVSEP ; VHATSEP $ ; SUBSE
*PRINT/P*     SET $ 8

```

```

*LABEL*       NEXT
*ENDMACRO*

```

```

*PRINT*       ANALYSIS
*LABEL*       NEXT_SET
*USE/R*       ANALYSIS   $
*DEVALUE*
*CALC*       I = I + 1
*JUMP*       NEXT_SET * ( I .LE. NSETS )
*RUN*

```

1 Number of sets
5 7 Set number and number of subsets

1 15 137
1 15 144
1 20 169
1 20 169
1 20 166
1 28 192
1 28 193
1 28 194
1 50 244

.
.
DATA SET with data SUBSET S V
.
.

7 28 199
7 50 238
7 50 243
50 231

```

*EOD*
*CLOSE*
*STOP*

```

THE DATA SET PLUS RS = 1/S AND RV = 1/V FOR SET 5

SUB_SET	S	V	RS	RV
1	1.5	0.137	0.667	7.299
1	1.5	0.144	0.667	6.944
1	2.0	0.169	0.500	5.917
1	2.0	0.169	0.500	5.917
1	2.0	0.166	0.500	6.024
1	2.8	0.192	0.357	5.208
1	2.8	0.193	0.357	5.181
1	2.8	0.194	0.357	5.155
1	5.0	0.244	0.200	4.098
1	5.0	0.249	0.200	4.016
1	5.0	0.247	0.200	4.049
2	1.5	0.138	0.667	7.246
2	1.5	0.145	0.667	6.897
2	1.5	0.146	0.667	6.849
2	2.0	0.166	0.500	6.024
2	2.0	0.167	0.500	5.988
2	2.0	0.167	0.500	5.988
2	2.8	0.195	0.357	5.128
2	2.8	0.186	0.357	5.376
2	2.8	0.192	0.357	5.208
2	5.0	0.230	0.200	4.348
2	5.0	0.228	0.200	4.386
2	5.0	0.226	0.200	4.425
3	1.5	0.137	0.667	7.299
3	1.5	0.138	0.667	7.246
3	1.5	0.136	0.667	7.353
3	2.0	0.158	0.500	6.329
3	2.0	0.161	0.500	6.211
3	2.0	0.166	0.500	6.024
3	2.8	0.180	0.357	5.556
3	2.8	0.187	0.357	5.348
3	2.8	0.184	0.357	5.435
3	5.0	0.229	0.200	4.367
3	5.0	0.232	0.200	4.310
3	5.0	0.229	0.200	4.367
4	1.5	0.133	0.667	7.519
4	1.5	0.136	0.667	7.353
4	1.5	0.141	0.667	7.092
4	2.0	0.166	0.500	6.024
4	2.0	0.161	0.500	6.211
4	2.0	0.169	0.500	5.917
4	2.8	0.187	0.357	5.348
4	2.8	0.190	0.357	5.263
4	2.8	0.192	0.357	5.208
4	5.0	0.231	0.200	4.329
4	5.0	0.227	0.200	4.405
4	5.0	0.233	0.200	4.292
5	1.5	0.121	0.667	8.264
5	1.5	0.124	0.667	8.065
5	1.5	0.129	0.667	7.752
5	2.0	0.150	0.500	6.667
5	2.0	0.145	0.500	6.897
5	2.0	0.150	0.500	6.667
5	2.8	0.170	0.357	5.882
5	2.8	0.171	0.357	5.848
5	2.8	0.174	0.357	5.747
5	5.0	0.202	0.200	4.956
5	5.0	0.203	0.200	4.926
5	5.0	0.210	0.200	4.762
6	1.5	0.150	0.667	6.667
6	1.5	0.153	0.667	6.536
6	1.5	0.153	0.667	6.536
6	2.0	0.176	0.500	5.682
6	2.0	0.182	0.500	5.495
6	2.0	0.184	0.500	5.435
6	2.8	0.210	0.357	4.762
6	2.8	0.203	0.357	4.926
6	2.8	0.209	0.357	4.785
6	5.0	0.251	0.200	3.984
6	5.0	0.255	0.200	3.922
6	5.0	0.253	0.200	3.953
7	1.5	0.137	0.667	7.299
7	1.5	0.149	0.667	6.711
7	1.5	0.147	0.667	6.803
7	2.0	0.171	0.500	5.848
7	2.0	0.174	0.500	5.747
7	2.0	0.174	0.500	5.747
7	2.8	0.198	0.357	5.051
7	2.8	0.190	0.357	5.263
7	2.8	0.199	0.357	5.025
7	5.0	0.238	0.200	4.282
7	5.0	0.243	0.200	4.115
7	5.0	0.231	0.200	4.329

DATA PLOT : INITIAL VELOCITY OF REACTION V
AGAINST SUBSTRATE CONCENTRATION S FOR SET 5

Reaction Velocity (V)	Substrate Concentration (S)	Notes	Point ID
0.2520			06
0.2505			06
0.2490			01
0.2475			01
0.2460			
0.2445			01
0.2430			07
0.2415			
0.2400			
0.2385			07
0.2370			
0.2355			
0.2340			
0.2325			03
0.2310			04
0.2295			02
0.2280			02
0.2265			02
0.2250			
0.2235			
0.2220			
0.2205			
0.2190			
0.2175			
0.2160			
0.2145			
0.2130			
0.2115			
0.2100	06		05
0.2085	06		
0.2070			
0.2055			
0.2040			
0.2025	06		05
0.2010			
0.1995	07		
0.1980	07		
0.1965			
0.1950	02		
0.1935	01		
0.1920	01		
0.1905	04		
0.1890			
0.1875	03		
0.1860	02		
0.1845	06		
0.1830			
0.1815	06		
0.1800			
0.1785	03		
0.1770			
0.1755	06		
0.1740	07		
0.1725			
0.1710	07		
0.1695	01		
0.1680			
0.1665	01		
0.1650			
0.1635			
0.1620			
0.1605	03		
0.1590			
0.1575	03		
0.1560			
0.1545			
0.1530	06		
0.1515			
0.1500	06	05	POINTS COINCIDING WITH POINT 1 2 3 4 7
0.1485	07		
0.1470	07		
0.1455	02	05	POINTS COINCIDING WITH POINT 2 2 3 4
0.1440	01		
0.1425			
0.1410	04		POINTS COINCIDING WITH POINT 3 4
0.1395			
0.1380	02		POINTS COINCIDING WITH POINT 4 7
0.1365	01		
0.1350			
0.1335	04		POINTS COINCIDING WITH POINT 5 5
0.1320			
0.1305			
0.1290	05		POINTS COINCIDING WITH POINT 6 6
0.1275			
0.1260			
0.1245	05		POINTS COINCIDING WITH POINT 7 7
0.1230			
0.1215	05		
0.1200			

1.20 1.60 2.00 2.40 2.80 3.20 3.60 4.00 4.40 4.80 5.20

S SUBSTRATE CONCENTRATION

V = RECIPROCAL (SLOPE * RS + INTERCEPT) + ERROR

285.....

***** REGRESSION ANALYSIS *****

ERROR DISTRIBUTION: NORMAL LINK FUNCTION: INVERSE

*** REGRESSION COEFFICIENTS ***

Y-VARIATE: V

	ESTIMATE	S.E.	T
CONSTANT	3.0951	0.0799	38.74
RS	5.9900	0.2322	25.80

*** ANALYSIS OF DEVIANCE ***

Y-VARIATE: V

TERMS	RESIDUAL DF	DEVIANCE	CHANGE DF	DEVIANCE	MEAN CHANGE	MN DEV. RATIO
INITIAL MODEL	82	*	*	*		
CONSTANT						

MODIFICATIONS TO MODEL

	82	0.1070135	0	*		
+RS	81	0.0103412	1	0.0966723	0.0966723	757.21
** DENOMINATOR OF RATIO IS (RES. DEV./RES. DF) FROM LINE ABOVE, = 0.0901277						

CALCULATE VMAX AND KM ESTIMATES FOR THE SET

SUMMARY

SET	5		
VMAXI	0.323086080	KM	1.935290112

ANALYSIS FOR EACH SUBSET	
SET	5

285.....

***** REGRESSION ANALYSIS *****
 ERROR DISTRIBUTION: NORMAL LINK FUNCTION: INVERSE

*** REGRESSION COEFFICIENTS ***

Y-VARIATE: V	ESTIMATE	S.E.	T
SUBSETS 1	2.7617	0.0660	41.86
SUBSETS 2	3.2793	0.0691	47.46
SUBSETS 3	3.1049	0.0708	43.85
SUBSETS 4	3.1100	0.0692	44.97
SUBSETS 5	3.5565	0.0861	41.32
SUBSETS 6	2.8509	0.0574	49.63
SUBSETS 7	3.0903	0.0646	47.84
RS.SUBSETS 1	6.5417	0.2046	31.97
RS.SUBSETS 2	5.5032	0.1948	28.24
RS.SUBSETS 3	6.3081	0.2066	30.53
RS.SUBSETS 4	6.0908	0.2004	30.39
RS.SUBSETS 5	6.4971	0.2464	26.37
RS.SUBSETS 6	5.4947	0.1660	33.10
RS.SUBSETS 7	5.5973	0.1846	30.31

*** ANALYSIS OF DEVIANCE ***

TERMS	RESIDUAL DF	RESIDUAL DEVIANCE	CHANGE DF	CHANGE DEVIANCE	MEAN CHANGE	MN DEV. RATIO
INITIAL MODEL						
CONSTANT						
RS	81	1.034E -2	*	*		
MODIFICATIONS TO MODEL						
-CONSTANT						
+SUBSETS	75	1.363E -3	6	8.978E -3	1.496E -3	112.29
-RS						
+RS.SUBSETS	69	9.194E -4	6	4.437E -4	7.395E -5	5.55
** DENOMINATOR OF RATIO IS (RES. DEV./RES. DF) FROM LINE ABOVE, = 1.333E -5						

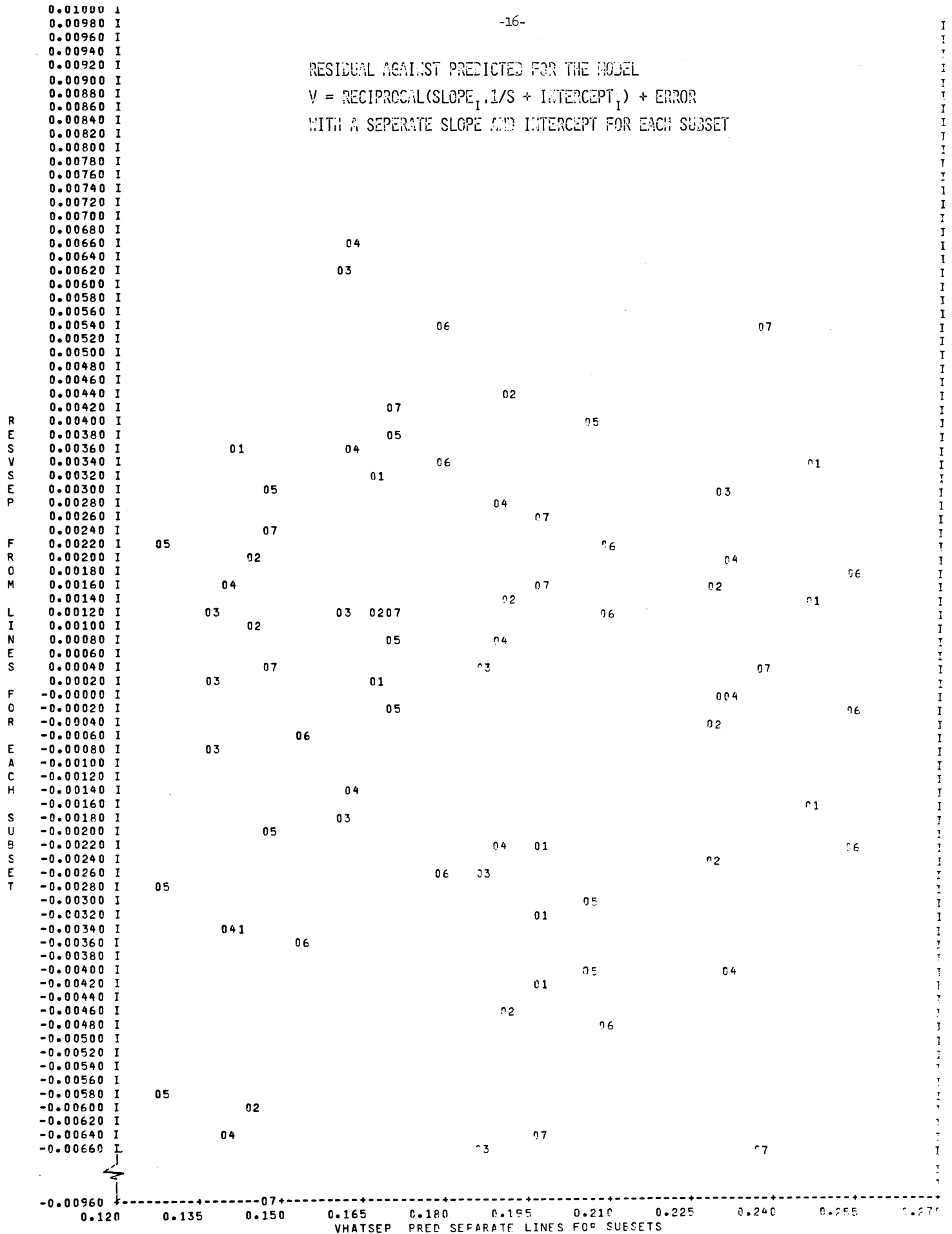
SUMMARY

SET 5
 VMAXI 0.323086080 KM 1.935290112

VMAX AND KM ESTIMATES FOR EACH SUBSET

SUBSET		VMAXI		KM	
SUBSET 1		0.362093568		2.368716032	
SUBSET 2		0.304944128		1.678170112	
SUBSET 3		0.322074624		2.031673424	
SUBSET 4		0.321545728		1.958476032	
SUBSET 5		0.281176064		1.826829824	
SUBSET 6		0.350763520		1.927339520	
SUBSET 7		0.323590912		1.811225856	

RESIDUAL AGAINST PREDICTED FOR THE MODEL
V = RECIPROCAL(SLOPE₁.1/S + INTERCEPT₁) + ERROR
WITH A SEPERATE SLOPE AND INTERCEPT FOR EACH SUBSET



Acknowledgment

Thanks go to Pat Firey for useful discussion.

References

- Alvey, N.G. et al. [1977]. GENSTAT. A general statistical program version 4.01. Statistics Department, Rothamsted Experimental Station.
- Cleland, W.W. [1970]. Steady state kinetics. The Enzymes. (P.D. Boyer ed.) 3rd edition, vol 2 pp 1-65, Academic Press, New York,
- Lineweaver, H. and Burk, D. [1934]. The determination of enzyme dissociation constants. J Am. Chem. Soc. 56: 658-666.
- Michaelis, L. and Menten, M. L. [1913]. Biochem Z. 49: 333.
- Nelder, J. A. [1966]. Inverse polynomials, a useful group of multi-factor response functions. Biometrics 27: 128-141.
- Nelder, J. A. and Wedderburn, R.W.M. [1972]. Generalized Linear Models. J. R. Statist. Soc. 135: 370-384.
- Walls, H. E. [1978]. Further studies of rat liver microsomal glucose-6-phosphatase. M. S. Thesis, Cornell University.