University of Southern California

School of Medicine

Laboratory of Applied Pharmacokinetics

Technical Report 99-1

NONLINEAR PARAMETRIC AND NONPARAMETRIC POPULATION PHARMACOKINETIC MODELING ON A SUPERCOMPUTER

By

Roger Jelliffe[1], Michael Van Guilder[1], Robert Leary[2], Alan Schumitzky[1], Xin Wang[1], and Alexander Vinks[3].

1 - Laboratory of Applied Pharmacokinetics, Univ of So Calif School of Medicine, 2250 Alcazar St (CSC-134-B) Los Angeles CA 90033

2 - San Diego Supercomputer Center, San Diego CA.

3 - The Hague Hospitals Central Pharmacy, The Hague, The Netherlands.

NONLINEAR PARAMETRIC AND NONPARAMETRIC POPULATION PHARMACOKINETIC MODELING ON A SUPERCOMPUTER.

Roger Jelliffe[1], Michael Van Guilder[1], Robert Leary[2], Alan Schumitzky[1], Xin Wang[1], and Alexander Vinks[3]. 1 - Laboratory of Applied Pharmacokinetics, Univ of So Calif School of Medicine, 2250 Alcazar St (CSC-134-B) Los Angeles CA 90033. 2 - San Diego Supercomputer Center, San Diego CA. 3 - The Hague Hospitals Central Pharmacy, The Hague, The Netherlands. Email jelliffe@hsc.usc.edu

KEYWORDS Population PK/PD Modeling, Parametric, Nonparametric, Pharmacokinetics, Supercomputer

ABSTRACT

We extended our nonparametric EM (NPEM) PC software to include nonlinear systems, incorporating a differential equation (DE) solver and parallelization. The user either draws the model with the BOXES program or enters the DE's on the PC, and runs another PC program to enter instructions for the SDSC Cray T3E, and to specify subject data files.

The user then uses SSH to access the Cray, gets the files from his local FTP node, compiles the DE source code, and runs the program. Speedup is 24-fold for 32 nodes. Using 8 nodes, a 5-parameter Michaelis-Menten NPEM model of pipericillin, using 80,000 grid points, took 1/2 hr. The user FTP's the result files to his PC and examines them with local software and plots.

INTRODUCTION AND OVERVIEW

Pharmacokinetic population analysis is used to quantify the intersubject variability in pharmacokinetic studies. It is a necessary tool in analyzing data from clinical trials. It provides the foundation for the intelligent design of drug dosage regimens.

Assume that each subject's drug response is a model which depends on the dosage regimen, the sampling schedule, measured covariates; and on the unknown model and noise parameters. The parameters are random variables with a common but unknown population distribution. Population analysis is the estimation of this population distribution based on the population data.

There are basically two approaches: parametric and nonparametric. In the <u>parametric</u> approach, the shape of the population distribution is assumed to be <u>known</u> except for the unknown population parameter values (e.g., normal with unknown means and covariances). The NONMEM system of Beal and Sheiner (1) is an example.

In the <u>nonparametric</u> approach, no such parametric assumptions about the form of the population parameter distributions are made. The entire distribution is estimated from the population data. It allows for non-normal and multimodal distributions such as occur, for example, in a population of fast and slow acetylators. This approach was developed by Mallet (3). A variation based on the nonparametric EM (NPEM) algorithm was developed by Schumitzky (4).

Both approaches use two methods for the optimal estimates and their variability: maximum likelihood and Bayesian. The Bayesian parametric method is described in (5). In the above terminology, the NPEM program is a maximum likelihood approach based on the nonparameteric EM algorithm. A good review of all approaches is found in (2).

EXAMPLE

Here we describe a Monte Carlo simulation study of the NPEM algorithm. We consider a

one compartment model with bolus input applied to a population of 20 simulated subjects.

The model for the ith subject is given by:
$$y_i = [\ 20*\exp(\ -K_i*t_j\ )/V_i\ ]\ (1 + e_j\ );$$
$j=1,...,M$, where

K_i = elimination rate constant

V_i = volume of distribution

M = # of measurements per subject

t_j = time of the jth measurement,

e_j = noise for j^{th} measurement.

It is assumed that the $\{e_j\}$ are independent and $e_j \sim N(0, 0.04)$ where the notation $x \sim N(m, s)$ means that the density of x is normal with mean = m and standard deviation = s. In the Monte Carlo simulation, the random vectors $(K_i,V_i)\}$ are independent and identically distributed as a mixture of two normals so K_i and V_i are independent and $V_i \sim N(2.0, 0..2)$, and $K_i \sim 0.5\ N(0.5, 0.05) + 0.5 * N(1.5, 0.15)$.

The true probability density of (K, V) is graphed in Figure 1. Note that in Figures 1-5, the "x-axis" corresponds to K; the "y-axis" corresponds to V.

Fig. 2. Normal density of a random vector with same mean and covariance as (K,V).

Two cases are considered: 5 and 2 levels per subject. Figure 3 shows the smoothed estimated density for the 5 levels case.

Figure 4 shows the smoothed estimated density (K, V) for the 2 levels case. There is not much deterioration from the loss of measurements. The densities are actually discrete.
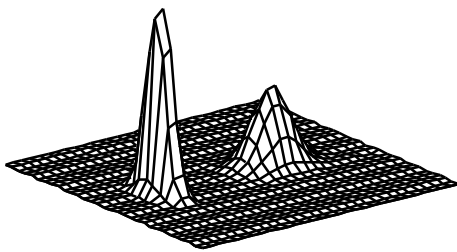

Fig. 1. True density of (K, V)

Note that if (K, V) were Gaussian, and were perceived by a parametric method, then the "best" estimate of this density would look like Figure 2, in which most of the mass is situated where in the true density there is no mass at all.
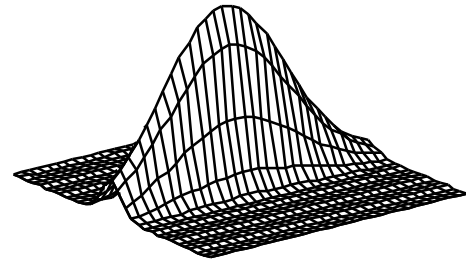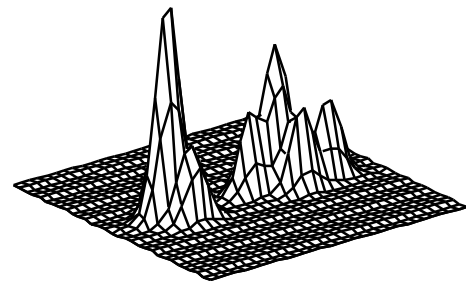

Fig. 3. Smooth estimated discrete density of (K, V), 5 levels / subject
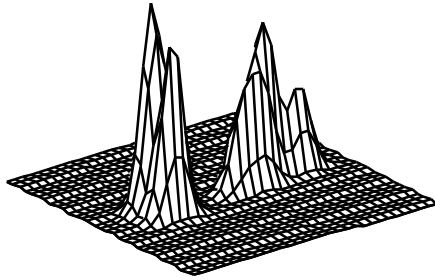
Fig. 4. Smooth estimated discrete density
of (K, V), 2 levels / subject

Figure 5 shows the similarly smoothed "empirical density" supported uniformly on the 20 actual points (K_i, V_i) used in the Monte Carlo simulation. It resembles Figures 3 and 4.
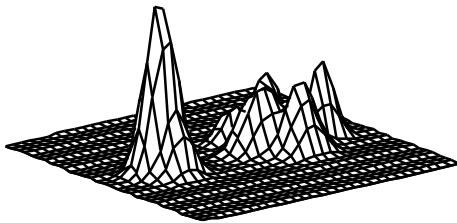

Fig. 5. Smooth empirical discrete density
of (K, V), 20 subjects

Thus the NPEM algorithm can discover previously unsuspected subpopulations of patients, while parametric methods cannot, without further help and further information. Furthermore, the NPEM method can do a good job of quantitatively perceiving the true discrete joint density.

PROCEDURE TO RUN 'BIG NPEM'

The "Big NPEM" now implemented on the SDSC Cray T3E has the following procedure for performing the analysis.

STEP 1. Run BOXES on the PC. Output=PROG.FOR, a Fortran file containing source code defining the model, or type in the Fortran source code. The structural model is shown in Figure 6 below.
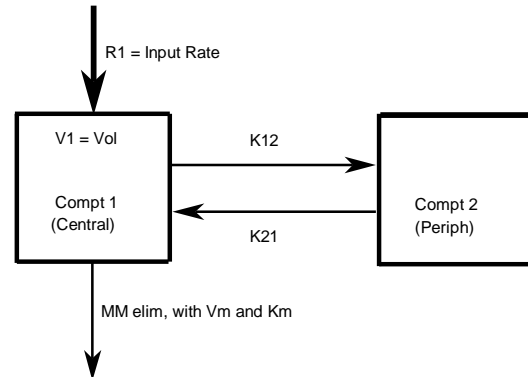

Fig 6. The Michaelis-Menten model of
pipericillin.

STEP 2. Run M2_20 on the PC, in DOS.
Input: Choose Option 1: Prepare files for 'BIG NPEM'
-PROG.FOR (Output from BOXES)
-Then one enters
        -Names and number of patient data files
        -Boundaries for random parameters
        -Assay Error Coefficients
        -Active or salt fraction
        -Number of grid points (from 2129 to about 20 million)
        -Max. number of cycles; stopping tolerance
        -AUC/predicted value selections
        -First density (uniform or previous density from a file)

Output: is 2 files. One is mod01.f=PROG.FOR (the Fortran source code for the model). The other is fil01.inp: all the above instructions, and the concatenated patient data files

STEP 3. Connect to SDSC Cray T3E, with Windows and SSH. Login to golden.sdsc.edu. Upload mod01.f and fi101.inp (ascii mode).

Execute comp.sh (compiles mod01.f, links it to BIG NPEM executable program.
Execute 'BIG NPEM': and state the no. of parallel cpu's.
In background, may logout and login later; use qcat command to monitor progress.
In foreground--watch output. Download output files from run by FTP to the PC. Logout at end of run.

STEP 4. Run M2_20 on the PC
Input:-Choose Option 2: Examine files from 'BIG NPEM'

-Enter names of output and density files, OUT0122, DEN0122, for example
-Select any or all of the 7 options to examine and plot results.

RESULTS

A partial display of the output for a Michaelis-Menten model of pipericillin, with IV input and both a central and a peripheral compartment is given below.

CYCLE NO. 4
THE TRUE (NUMERICAL) LOG-LIKELIHOOD (USING THE NUMBER THEORETIC INTEGRATION SCHEME) OF THE 8 SUBJECT VECTORS,
GIVEN THE PRIOR DENSITY, IS:
-560.82148881339288

THE DIFFERENCE BETWEEN THE LIKELIHOOD OF THE MAXIMUM LIKELIHOOD ESTIMATE OF THE DENSITY AND THE LIKELIHOOD OF THIS CYCLE IS LESS THAN THE FOLLOWING NUMBER: 0.62775346607536164

THE CORRESPONDING FIGURE FOR CYCLE 1 WAS 80014. SINCE IT IS NOW .62775 , THIS ANALYSIS HAS GONE 80013. DIVIDED BY 80014. OR 99.99922 % OF THE WAY FROM THE APRIORI DENSITY TO THE MAXIMUM LIKELIHOOD ESTIMATE OF THE JOINT DENSITY.
THE NO. OF ACTIVE GRID POINTS IS NOW 29

THE FOLLOWING VALUES ARE FOR THE UPDATED DENSITY:

THE SCALED INFO FOR THIS CYCLE IS 98.91 %
THE ENTROPY FOR THIS CYCLE IS 2.1796

THE MEANS ARE:

VM MM 1-	KM MM 1-	V1
2195.214891	85.541580	13.224692
K12	K21	
0.802670	0.57	

THE STANDARD DEVIATIONS ARE, RESPECTIVELY:

VM MM 1-	KM MM 1-	V1
818.093945	60.553616	3.840690
K12	K21	
0.627348	0.44	

THE PERCENT COEFFICIENTS OF VARIATION ARE, .:

VM MM 1-	KM MM 1-	V1
37.267146	70.788518	29.041810
K12	K21	
78.157675	76.54	

Further results are shown graphically in the figures below.

APPROXIMATE MARGINAL DENSITY

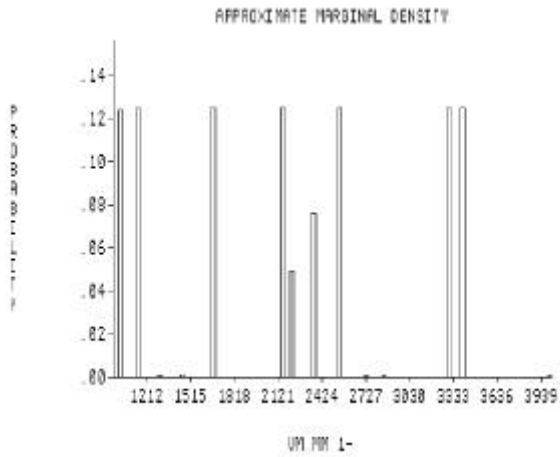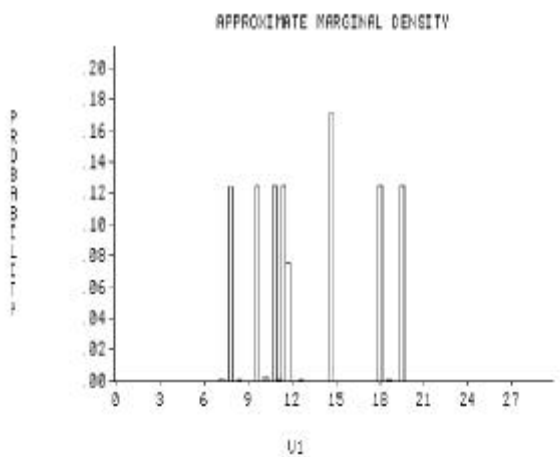Fig 7. Discrete Marginal Density of Vm

APPROXIMATE MARGINAL DENSITY

Fig 8. Discrete Marginal Density of Km
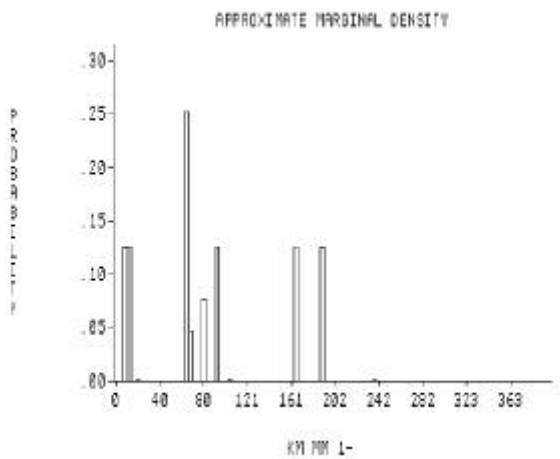
APPROXIMATE MARGINAL DENSITY

Fig 9. Discrete Marginal Density of V1 (Vol)

APPROXIMATE MARGINAL DENSITY

Fig 10. Discrete Marginal Density of K12

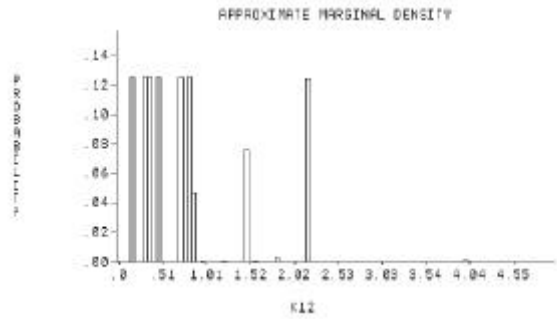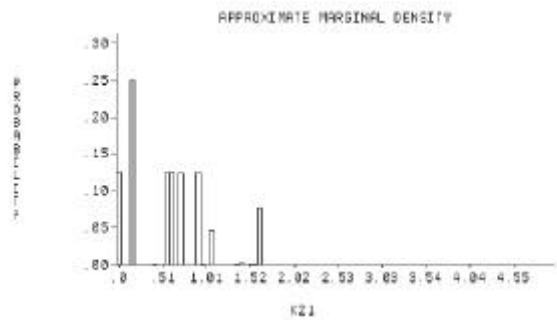APPROXIMATE MARGINAL DENSITY

Fig 11. Discrete Marginal Density of K21

Further graphics illustrate the correlation relationships between parameter pairs, as shown below. These graphics can be compared with the correlation coefficient matrix (not shown here, to save space).

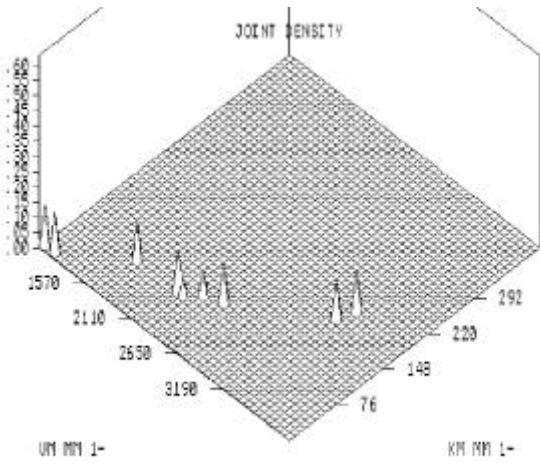The strong correlation between Vm and Km is shown in Figure 12.

Fig 12. Discrete Marginal Joint Density of
Vm and Km

The relationship between K12 and K21 is shown below.
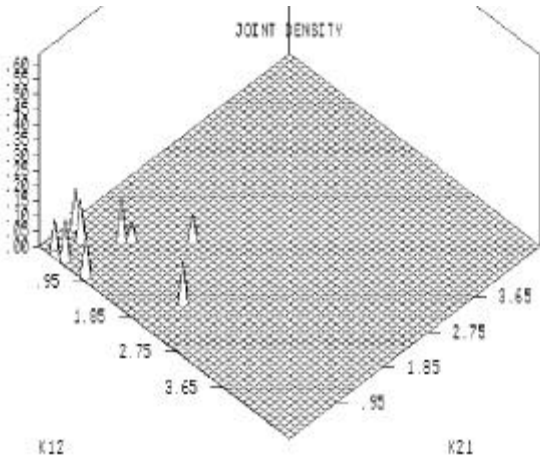

Fig 13. Discrete Marginal Joint Density of
K12 and K21

The scattergram of concentrations estimated using the population median parameter values is shown below.
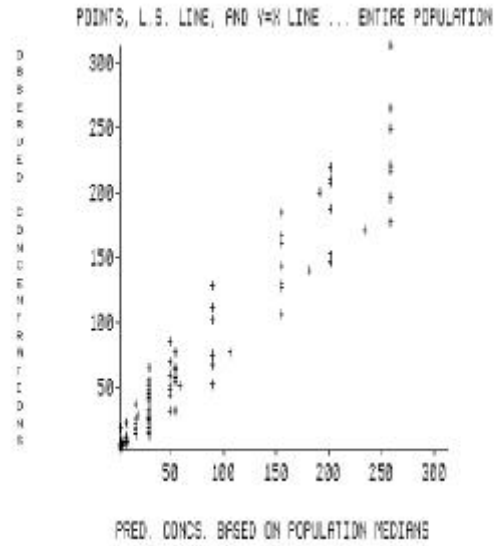

Fig 14. Scattergram of estimated serum concentrations based of population parameter median values.

The similar scattergram of serum concentrations, this time estimated from the joint Bayesian posterior parameter medians is shown below.
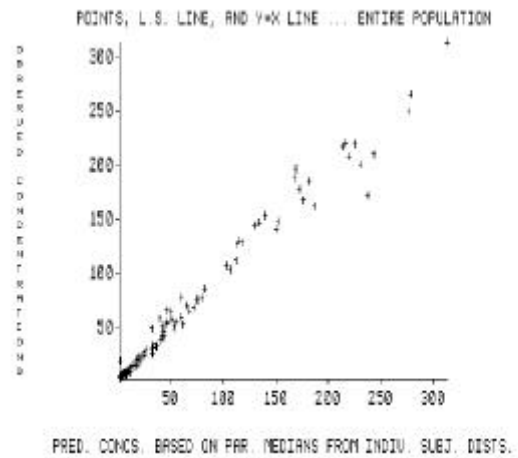

Fig 15. Scattergram of estimated serum concentrations based of individual subject Bayesian posterior median values.

This software has now been extended to include the parametric iterative Bayesian Two-Stage (IT2B) algorithm as well.

The combination of these two methods for making large and nonlinear population PK/PD models is a strong one. The IT2B approach permits separation of variability associated with the assay error from that associated with the subject intraindividual variability. Parameter ranges are also easily estimated.

The above information is then usefully incorporated into the NPEM approach. The strength of the nonparametric approach is then that it leads directly to the new "multiple model" (MM) design of dosage regimens which are specifically designed to achieve desired target goals with maximal precision (minimal weighted squared error of the diverse predicted serum concentrations, approximately one per population subject) when the population model is used as the Bayesian prior for designing the initial dosage regimen (6).

Because of this, nonparametric population models and MM dosage regimens are useful for developing the optimal standard dosage regimen, especially when it must be used without feedback from measured serum concentrations. Whether one is searching for a standard dose for a population of subjects, as in cancer chemotherapy or therapy for AIDS, or whether one must treat a herd of cattle or flock of chickens, for example, the combination of a nonparametric population model and MM dosage design assures maximally precise achievement of desired target goals, whether they be serum concentrations or quantifiable effects.

The NPEM software has now been upgraded to include up to 10 covariates or descriptors of categorical or numerical types, and up to six separate responses (serum concentrations or effects). Multiple inputs, to permit modeling of combination drug therapy and its various shared and separate therapeutic and/or toxic effects are now being implemented.

The above IT2B and NPEM software has now become a research resource for such modeling of systems. The grid points have been increased from 80,000 on the PC to about 20 million to permit the larger models.

REFERENCES
1. Beal, S., and L. Sheiner. (1992) *NONMEM User's Guide.* NONMEM Project Group, University of California, San Francisco.

2. Davidian M. and D. Giltinan (1995). *Nonlinear Models for Repeated Measurement Data.* Chapman and Hall, New York.

3. Mallet, A. (1986) A Maximum Likelihood Estimation Method for Random Coefficient Regression Models. *Biometrika* 73: 645-656.

4. Schumitzky, A. (1991) Nonparametric EM Algorithms for Estimating Prior Distributions. *Applied Mathematics and Computation* 45: 141-157.

5. Wakefield J., A.F.M. Smith, A. Racine-Poon and A. Gelfand. (1994) Bayesian Analysis of Linear and Nonlinear Population Models. *Applied Statistics* 43: 201-222.

6. Jelliffe R, Schumitzky A, Bayard D, Milman M, Van Guilder M, Wang X, Jiang F, Barbaut X, and Maire P: Model-Based, Goal-Oriented, Individualised Drug Therapy. Clin Pharmacokinet 34: 57-77, 1998.