TRACKS INDEX

Track 8 "IT Security & Maitenance"

PREDICTION OF SQL PROGRAM MAINTENANCE[1] EFFORT

Antonio Martínez
Research Group Alarcos, University Of Castilla-La Mancha, Ronda De Calatrava, 5, 13071, Ciudad Real (España)
E-Mail: Amartinez@Inf-Cr.Uclm.Es
excma. Diputacion De Ciudad Real
Calle Toledo, 17, 13001, Ciudad Real (Spain)

Mario Piattini
Research Group Alarcos, University Of Castilla-La Mancha, Ronda De Calatrava, 5, 13071, Ciudad Real (España)
E-Mail: Mpiattin@Inf-Cr.Uclm.Es

Abstract

Fourth Generation Language environments are substituting more and more Third Generation Language's (mainly COBOL), as a platform for computer system development. As a result, it is very important that software project managers are able to predict 4GL program maintenance effort. A way of carrying out this control is through the use of specific metrics for these environments. This is a field of software engineering where little research has been done.

Using simple regression analysis, a prediction model of implementation maintenance with an accuracy of MMRE=23% and pred (0,25)=86,71% was developed.

This model is based on different metrics of the SELECT statement which are easily computed.
Keywords: SQL, Metrics maintenance effort prediction, multivariate regression.

1. INTRODUCTION

Many organisations which use management information systems are now realising that computer systems developed using third generation languages such as COBOL can be more effectively produced and maintained using modern productivity-enhancing tools (Holloway,1990).

[1] This work is part of MANTICA project; it is financed by CICYT and the European Union (1FD97-0168)

Metrics[2] are useful mechanisms in improving the quality of software products especially maintenability. Maintenance is the most important problem in software development ranging between 60 and 90 percent of life-cycle costs (Card and Glass,1990; Pigoski,1997). Software measurement is widely recognised as an effective means to understand, monitor, control, predict and improve software development and maintenance projects (Briand et. al, 1996).

Different types of metrics have been defined for 4GL. So far, some projects have been developed to estimate development effort and the correlation of these with the size of a program (Dolado,1997; Verner and Tate,1988). Bourque and Côte describe an empirical study to predict 4GL system size based on several metrics derived in E/R diagrams (Bourque and Côte,1991). Using linear regression they were able to develop effective prediction systems although they noted the need to calibrate the models to the specific measurement environment. Ince et al. and Gray et al suggested a similar approach and indeed our data collection includes the raw counts required for the more complex synthetic metrics proposed by these authors (Ince et. al,1991;Gray et. al, 1997).

Niessink and Van Vliet propose to use maintenance Function Points to estimate the effort needed to implement change requests. This result is MMRE=71% and PRED(25)=21% for unadjusted maintenance function points (MFP) and MMRE=73% and PRED(25)=32% for adjusted maintenance function points (AMFP) (Niessink and Vliet,1997).

Jorgensen chose lines of code (LOC) and function points for the measurement of software maintenance task size. The measure "function points" was considered for size measure but had several disadvantages. For example, the number of function points is not very useful on small maintenance tasks, change oriented tasks and on maintenance tasks which are not concerned with user functionality (Jorgensen,1995).

A measure based on LOC has disadvantages, as well. The size in LOC does not reflect task characteristics very well in areas such as, change of software functionality, quality and usefulness. Neither does it reflect most of what the maintainer really does, such as test design, documentation, and analysis activities. Better measures on the size of a maintenance task are therefore needed, and research on this topic may be important in order to improve the quality of the maintenance effort prediction models.
This paper addresses development effort and not maintenance effort.
In MacDonell et al. an empirical study for building prediction systems is carried out. The authors want to predict the size of 4GL systems using a number of non-menu functions as a prediction with MMRE=21%. There are a number of reasons for avoiding function points, mainly due to complexity of collection and the subjectivity of the process (MacDonell et. al,1997).

---

[2] For present purposes, the words measure and metric are used interchangeably. For more precise definitions (Melton,1996).

However, the prediction systems implied by their work remain unvalidated. We are working to develop a set of metrics to predict the maintenance effort of fourth generation environment programs.

2. PROPOSED METRICS

We propose three kinds of metrics for the data base manipulation sub-language and singularised it to SELECT statement as follows:

NT measure

Number of tables referred to in the SELECT statement.

NN measure

Number of nesting, considers the number of "SELECT" in the SELECT statement.

NG measure

Number of GROUP BY in the SELECT statement.

This measure verifies the properties, nonnegativity, null value, module additivity of size metric (Briand et. al,1996;Martínez and Piattini,1998).

In the example of figure 1 the values are NT=3, NN=3 and NG=1.

```
select f.name_emp, p.number_fic, p.date
      from control_employee p, employee f, h_employee h
      where p.id_emp not in
            (select h.id_emp
           from control_employee p, employee f, h_employee h
             where p.number_fic=h.number_fic
                    and f.id_emp=h.id_emp
                     and p.id_emp=f.id_emp
                     and p.date='171298'
                     and p.control='SM'
                    and p.status='A'
                     and f.sex='V'
                    and p.hour in
            (select hour
                  from control_employee p, employee f, h_employee h
                  where p.number_fic=h.number_fic
                                and f.id_emp=h.id_emp
                                and p.id_emp=f.id_emp
                                and p.date='171298'
                                and p.control='SM'
                                and p.tipe='A0'
                                and h.remainder=0
                        )
            )
            and p.date='151298'
            and p.number_fic=h.number_fic
            and p.id_emp=f.id_emp
            and f.id_emp=h.id_emp
 group by f.name_emp, p.number_fic, p.date
```

Figure 1. Example of a SELECT statement

These measures were proposed based on intuition and experience with SQL programs development and maintenance. The number of tables are likely to influence all the three maintainability factors: understandability, modifiability and testability (Li and Cheng,1987). This is due to the fact that SQL statements are more difficult to understand, modify and to test if they include more tables.

The number of nesting is also likely to influence the maintainability of the SQL code as each nesting demands a new level of thinking similar to a new call in 3GL or a level of inheritance in object-oriented programs (Cant et. al, 1995). Earlier relational optimizers had also been influenced by the SELECT nesting, and vendors recommended not to nest beyond 3 levels for performance reasons.

Grouping rows for calculating values also should influence the maintainability of SQL programs as it implies an additional operation which must be carried out over a set of rows.

3. STUDY CASE

3.1. General System Characteristics

The system is composed of 143 programs developed over a period of one year at the Data Processing Center of the Ciudad Real County Council. The system is a transaction processing system for data maintenance. The programs with embedded SQL are all of small to medium size: each SELECT statement included an average of six tables and two-level nesting (table 1).

More importantly, the system developed was functionally sound, providing an actual working solution to an actual organizational problem.

One of the most positive aspects of the system is the fact that it was constructed completely by one programmer, employing the same methodology and the same developing environment the CA-OpenIngres/4GL. These common factors are advantageous in they can be considered as constants in the analysis, a condition not often encountered in software size research. When they vary, factors such as these can have an obvious impact on system size. Given that these potential contributors may be treated as constant, the degree of confidence adopted in regard to any size relationships supported by the data will consequently be greater (MacDonell et. al,1997).

3.2 Data Collection

The maintenance process includes adding functionality to the software (adaptive maintenance) and correcting defects discovered in the systems (corrective maintenance).

Our study examined data from a maintenance period beginning with the original installation of the product and ending the product's second release. The required changes varied in magnitude from a simple command line option change to a more complex one. During the maintenance period the programmer recorded daily effort maintenance product, descriptions of the faults encountered as a consequence of enhancement activities, time spent correcting these faults and the functions to which the time and faults could be attributed. This tool is implemented in CA-OpenIngres.

3.3 Data Analysis and Results

3.3.1 Descriptive Statistics

General descriptive statistics for each one of the variables are shown in table 1.

Table 1: Descriptive statistics for each measure

Variable	Mean	Variance	S.E. Skew	Min	Max	Std Dev	Skewness
NT	6,042	30,364	0,203	0	22	5,510	1,361
NN	1,923	3,438	0,203	0	6	1,854	0,712
NG	0,399	0,241	0,203	0	1	0,492	0,419
TIME	74,364	5.838,865	0,203	1	290	76,413	1,205

Valid observations - 143 Missing observations – 0

3.3.2 Correlation Analysis

For the correlation test we use Pearson´s coefficient statistics and Spearman´s non-parametric correlation. We do this in order identify potential relationships between the variable time of maintenance effort (expressed in minutes) and the measures defined, as well as the relationships that could exist between the same variables. The results are shown in tables 2 and 3.

The statistics of both correlation sets show strong significant relationships between the variable time of maintenance and the defined measurement, except for NG measurement. We can observe that the relationships between the specification of the measured NT, NN and the variable TIME of maintenance are significant.

Table 2. Pearson's Correlation Coefficients

		NT	NN	NG	TIME
Pearson	NT	1,000	0,796	0,280	0,986
	NN	0,796	1,000	0,258	0,881
	NG	0,280	0,258	1,000	0,284
	TIME	0,986	0,881	0,284	1,000
Sig	NT		0,000	0,001	0,000
	NN	0		0,002	0,000
	NG	0,001	0,002		0,001
	TIME	0	0,000	0.001	
N	NT	143	143	143	143
	NN	143	143	143	143
	NG	143	143	143	143
	TIME	143	143	143	143

Table 3. Spearman's correlation coefficients.

		NT	NN	NG	TIME
Spearman	NT	1,000	0,799	0,286	0,964
	NN	0,799	1,000	0,243	0,908
	NG	0,286	0,243	1,000	0,283
	TIME	0,964	0,908	0,283	1,000
Sig	NT		0,000	0,001	0,000
	NN	0,000		0,003	0,000
	NG	0,001	0,003		0,001
	TIME	0,000	0,000	0.001	
N	NT	143	143	143	143
	NN	143	143	143	143
	NG	143	143	143	143
	TIME	143	143	143	143

NT and NN are also highly correlated which is logical because each nesting introduces a table. This is normally different from the table of the previous nesting level. We are conscious that maintenance effort can depend on several different factors other than the SELECT characteristics. Some programs have, besides SELECT, other statements like INSERT, DELETE or UPDATE, as well as different procedural and visual statements. However, due to the type of the programs involved, we think that these results are a valid first attempt to characterize SQL programs.

The level of grouping (NG) is not correlated with the time of maintenance. We are unable to find an answer for this at the moment. Perhaps the low occurrences of grouping in the programs in the case study might explain this low correlation. However more case studies and experiments must be considered in order to explain the influence of grouping in SQL understandability, modifiability and testability.

3.3.3. Regression analysis

A strong correlation between the two variables (NT, NN) defined in lineal multiple regression model does exists. It does so in order to determine the maintenance time from the specification of the measurement, illustrated by their values in Pearson´s correlation statistic (see table 2).

The lineal multiple regression model is utilized stepwise selecting an application developed in 4GL consisting of 143 programs (observations) giving the following model of regression (see table 4):

$$TIME = -11.513 + 10.789(NT) + 10.758(NN)$$

Table 4. Coefficients Coefficients[a]

Model		Not standardized Coefficients		Standardized Coefficients	T	Sig.
		B	Error típ.	Beta		
1	(Constant)	-8,243	1,593		-5,174	,000
	NT	13,672	,195	,986	70,065	,000
2	(Constant)	-11,513	,533		-21,593	,000
	NT	10,789	,106	,778	101,571	,000
	NN	10,758	,316	,261	34,075	,000

a. Dependent variable : TIME

This model has an R^2 adjusted to 0.997, indicating that 99% of the variance in implementation of variable time maintenance. The value of R^2 gives a measurement of the consistency in a specific model of regression. This was a particularly pleasing result suggesting that a significant degree of maintenance time could be attributed (or at least predicted) to end of application.

A question may be raised as to the increase in acceptability of the model, given the inclusion of two independent terms which, in fact, are related to each other as illustrated in the previous correlation table (at 0,796 Pearson and 0,799 Spearman). A model of lineal regression was also carried out including only the most significant independent variables. The form of this model is shown next:

TIME = -8.243 + 13.672(NT)

This model has an associated R^2 adjusted to 0.972, a lesser value compared with the model with two variables. A technique that could help to determine the gain associated with the inclusion of extra terms in a regression model is the R^2 test adequate (Neter et. al, 1983) :

$$R^2_{sub} > 1 - (1 - R^2_{full}) (1 + d_{n,k})$$

where:

R^2_{sub} is the value obtained of R^2 with the subset of predictor variables.

R^2_{full} is the value obtained of R^2 with all the predictor variables.

$d_{n,k} = (k F_{k,n-k-1}) / n - k -1$

In this case:

$R^2_{full} = 0.997$ (to the model of two predictor variables) (see table 5)

$R^2_{sub} = 0.972$ (to the model of a predictor variable) (see table 5)

$R^2_{sub} > 0.968$

Since the value of R^2 in the model with only one predictor variable (0.972) is higher than the minimum threshold of the value of adequate R^2 for the complete model with two predictor variables, (0.968). We could say that the model that includes the term NT alone is as effective in terms of consistency as the model containing both variables, NT and NN.

An important and desirable aspect of the regression model is the consistency illustrated by R^2 although this is not enough. As a matter of fact, a model might be very consistent, but excessively inaccurate in terms of correspondence to individual values pairs. Indicators commonly used in software metrics data analysis evaluate the accuracy of regression models are the mean magnitude of relative error (MMRE) and the threshold-oriented pred[3] measure.

The magnitude of relative error (MRE) is a normalised measurement of the discrepancy between current values (V_A) and estimated values (V_F):

$$MRE = Abs\ (V_A - V_F) / V_A)$$

The mean MRE (MMRE) is therefore the mean values for this indicator with all observations in the sample. A lower value for MMRE generally indicates a more accurate model.

The Pred measure provides an indication of overall fit for a set of data points, based on the MRE values for each data pair:

$$Pred\ (l) = i / n$$

Where:
l is the threshold value selected for MRE.
i is the number of pairs of data with MRE less than or the same as *l*.
n is the total number of pairs of data.

In this case study, MMRE=0,236870 and pred(0,25) = 119/143=0,867, then we could say that a 86,7% of the values estimated falls within the 25% of their corresponding actual value.

3.3.4. Residual analysis

Residual analysis, in which predictive errors are considered in relation to both the estimated and actual values indicated a slight tendency to overestimate for smaller programs. However, residual plot examination did not show any significantly problematic trends in error distribution. (Durbin-Watson = 2,037 see table 5).

Table 5. Residual analysis

Model	R	R Square	Adjusted R Square	Standard Error	Durbin-Watson
1	,986[a]	,972	,972	12,81	
2	,998[b]	,997	,997	4,22	2,037

[3] *Prediction at level l (Pred(l)),* where *l* is a percentage, is defined as the quotient of number cases in which the estimates are within the *l* absolute limit of the actual values divided by the total number of cases

3.3.5. Outliers (extreme significant observations).

A number of outlying data points were encountered in the data set. As a result, these points were examined in greater depth to determine whether the observations were valid within the context of the sample. In particular, some observations stood out as significant extreme outliers with an associate MRE of 3.30 (appendix B) under the model of the two variables. This value was met in the programs that had a number of tables and nestings approximate to zero, but associated maintenance time.

4. CONCLUSIONS

An important task for any software project manager is to be able to predict and control project maintenance effort. Unfortunately, there is comparatively little work, other than function points, that tackles the problem of building prediction systems for software that is dominated by data considerations. In particular systems developed using 4GLs (MacDonell et. al, 1997).

As for the proposed metric validation in the real case study and its results, we can say that the metrics affect the effort maintenance time of the SQL programs. In particular, in regression analysis we saw that number of the tables (NT) and the number of nesting (NN) are the metrics which affect to a greater extent the programs maintainability.

In this study we have correlated the measure lines of code (LOC) with the measure number of select (NS). The result was obtained from MMRE=288,686% and PRED(0,25)=14,68%. Our results are better (MMRE=21% and PRED(0,25)=86,71% compared with MMRE=288,686% and PRED(0,25)=14,68%.

This study has some obvious limitations in as one programmer was involved in the study. To corroborate futher the results, a larger study with more programmers is needed.
We are investigating the generalisation of other 4GL applications.

We are studying other types of maintenance tasks (perfective or preventative).

Use others approaches different methods of statistical analysis by linear regresion for prediction models as genetic programming and neural networks (Dolado, et. al, 1999).

5. REFERENCES

Bourque, P. y Côté, V. (1991). An experiment in software sizing with structured analysis metrics. *Journal of Systems and Software 15*, 159-172.

Briand, L.C., Morasca, S. y Basili, V. (1996). Property-based software engineering measurement. *IEEE Transactions on Software Engineering,* 22(1), 68-85.

Cant, S. N., Jeffery, D.R., and Henderson-Sellers, B, (1995). A concept model of cognitive complexity of element of the programming process. *Information and Software Technology*. 37(7), 351-362.

Card, D.N. y Glass, R.L. (1990). *Measuring Software Design Quality*. Englewood Cliffs. USA.

Dolado, J.J. (1997). A Study of the Relationships among Albrecht and Mark II Function Points, Lines of Code 4GL and Effort. *J. Systems Software,* 37,161-173.

Dolado, J.J. , Fernandez, L., Otero, MC. y Urkola, L. (1999). Software effort estimation: the elusive goal in project management. *ICEIS,* 412-418.

Gray, R.H.M., Carey, B.N., McGlynn, N.A. and Pengelly, A.D. (1997) 'Design metrics for database systems'. *BT Technology Journal,* 9, 4, 69-79.

Holloway, S. (ed.) (1990). *Fourth-Generation Systems, their scope application and methods of evaluation*. London: Chapman and Hall.

Ince, D.C., Shepperd. M.J., Pengelly, A. And Benwood,H. (1991). 'The metrification of data designs', in *Proc 3rd Annual Oregon Workshop on Software Metrics*, March 17-91.

Jorgensen, M (1995), Experience With the Accuracy of Software Maintenance Task Effort Prediction Models. *IEEE Software* 674-680.

Li, H.F. and Cheng (1987), W.K, An empirical study of software measures. IEEE Trans. on Software Engineering, 13 (6), 679-708.

Martínez, A. y Piattini, M. (1998). Validation of 4GL metrics. *Proc. of the Software Measurement in Practice, 10th Anniversary Conference*. United Kingdom Software Metrics Association, Londres, octubre 1998, X 1-19.

MacDonell, G., Shepperd y J., Sallis, J. (1997). Metrics for Database Systems: An Empirical Study. *IEEE Software* 99-107

Neter,J., Wasserman, W. y Kutner, M.H. *Applied Linear Regression Models*. Irwin: Homewood IL, 1983.

Niessink, F. and Vliet, H (1997), Predicting Maintenance Effort with Funcition Points. *IEEE Software* 32-39.

Pigoski, T.M. (1997). *Practical Software Maintenance*. Wiley Computer Publishing. New York, USA.

Verner J. y Tate G.(1988) Estimating Size and Effort in Fourth-Generation Development. *IEEE Trans. On Software Engineering*, 15-22.

Zuse, H. (1998). *A Framework of Software Measurement* , Ed. Walter De Gruyter.

APPENDIX A

The set data is presented below in the following order: Name (name of SQL programs), NT (number of tables), NN (number of nesting), NG (number of grouping), TIME (time in minutes).

Name	NT	NN	NG	TIME
Timer_on.osq	1	0	0	2
compr_li.osq	5	2	0	65
consulta.osq	10	3	1	130
cont000.osq	3	1	0	30
cont100.osq	3	1	0	30
cont101.osq	8	2	1	95
cont102.osq	6	2	1	75
cont103.osq	12	4	1	160
cont104.osq	3	1	0	30
cont105.osq	4	0	1	30
cont110.osq	1	0	0	3
cont120.osq	4	3	0	65
cont121.osq	7	4	1	105
cont140.osq	10	3	1	130
cont141.osq	3	2	0	40
cont200.osq	2	3	0	45
cont201.osq	3	0	1	20
cont202.osq	4	2	0	40
cont203.osq	5	3	1	75
cont204.osq	3	2	0	40
cont205.osq	14	5	1	195
cont205a.osq	5	1	0	50
cont206.osq	20	4	0	250
cont207.osq	20	5	1	260
cont208.osq	22	6	0	290
cont209.osq	22	6	1	290
cont240.osq	18	4	1	225
cont241.osq	5	3	0	75
cont242.osq	6	2	1	75
cont299.osq	1	0	0	3
cont300.osq	4	0	0	45
cont301.osq	1	0	0	3
cont302.osq	1	0	1	3
cont303.osq	5	3	0	70
cont304.osq	3	1	0	30
cont305.osq	4	0	0	30
cont306.osq	7	3	0	95
cont310.osq	10	4	1	135
cont311.osq	15	6	1	225
cont312.osq	1	0	0	3

Cont. APPENDIX A...

Name	NT	NN	NG	TIME
cont315.osq	5	1	1	30
cont316.osq	7	2	0	90
cont320.osq	11	3	0	140
Cont321.osq	20	3	1	250
cont321d.osq	18	2	1	200
cont322.osq	2	1	1	30
cont323.osq	9	2	1	105
cont330.osq	5	0	0	40
cont331.osq	4	1	0	40
cont332.osq	3	1	0	30
cont335.osq	2	1	1	20
cont336.osq	10	4	0	135
cont340.osq	12	3	1	150
cont341.osq	10	5	0	150
cont342.osq	5	2	1	65
cont343.osq	4	1	0	40
cont344.osq	1	0	0	10
cont345.osq	3	0	0	20
cont346.osq	3	1	0	30
cont347.osq	4	0	1	30
cont348.osq	4	1	0	40
cont361.osq	4	2	0	50
cont362.osq	4	3	0	65
cont400.osq	1	0	0	3
borra_u.osq	5	3	1	75
cont401.osq	4	0	0	30
cont402.osq	6	3	0	85
cont403.osq	15	6	1	215
cont410.osq	3	0	0	20
cont411.osq	2	1	0	20
cont412.osq	2	1	0	20
cont413.osq	4	2	1	55
cont414.osq	2	1	0	10
cont500.osq	4	1	1	40
cont501.osq	2	0	1	10
cont502.osq	4	2	1	50
cont503.osq	1	0	0	3
cont510.osq	3	2	1	45
cont511.osq	10	3	0	130
cont512.osq	3	2	0	40

Cont. APPENDIX A...

Name	NT	NN	NG	TIME
cont513.osq	16	3	0	190
cont515.osq	18	4	1	225
cont521.osq	13	5	0	180
cont524.osq	13	4	0	170
cont525.osq	12	5	1	175
cont600.osq	3	0	1	20
cont601.osq	6	4	0	95
cont602.osq	5	2	0	60
cont604.osq	4	1	1	40
Cont610.osq	1	0	1	1
Cont611.osq	3	0	1	20
Cont612.osq	2	1	1	25
Cont613.osq	1	0	1	3
Cont614.osq	3	0	0	15
Cont615.osq	9	6	0	150
Cont616.osq	1	0	0	1
Cont620.osq	6	3	0	90
cont621.osq	5	3	0	75
cont640.osq	1	0	1	20
cont641.osq	0	0	0	5
cont700.osq	1	0	0	10
cont701.osq	4	5	1	90
cont702.osq	3	0	0	20
cont705.osq	1	0	0	3
cont706.osq	2	0	0	10
cont997.osq	1	0	0	3
cont998.osq	2	0	0	10
cont999.osq	1	0	0	3
contu000.osq	3	1	1	65
conx501.osq	2	1	0	35
conx502.osq	3	0	0	20
conx503.osq	0	0	0	3
conx510.osq	5	2	0	60
conx511.osq	12	3	0	155
conx512.osq	4	2	1	50
conx513.osq	22	6	1	290
conx515.osq	19	4	0	240
conx521.osq	14	6	0	205
conx524.osq	13	5	1	180
conx525.osq	10	4	1	140

Cont. APPENDIX A...

Name	NT	NN	NG	TIME
conx526.osq	4	2	0	50
conx527.osq	14	5	1	190
dir_act.osq	4	0	1	30
edic_mor.osq	2	0	0	10
edic_tbl.osq	8	1	0	85
error1.osq	2	1	0	20
error1c.osq	3	0	1	20
fich_tab.osq	3	1	0	30
fich_app.osq	1	0	0	3
hay_dato.osq	2	0	0	10
importa.osq	2	0	1	10
inicial.osq	3	0	0	20
medir_fic.osq	3	0	0	20
modifica.osq	2	1	1	20
ponrev.osq	1	0	0	3
producti.osq	11	2	1	130
puesto.osq	3	0	1	15
reloj203.osq	3	0	0	20
reloj205.osq	17	6	1	235
reloj240.osq	17	4	0	215
reloj403.osq	10	6	1	160
sal_tbl1.osq	2	0	0	10
timer_o.osq	1	0	0	3

APPENDIX B

The result of the application the multiple lineal regression to the set of original data is present in this appendix:

Diagnostic per case(a)					
Number of case	**Tip. Residual**	**Time**	**Predicted value**	**Residual**	**MRE**
1	0,646	2	-0,72	2,72	1,36
2	0,249	65	63,95	1,05	0,016153846
3	0,319	130	128,65	1,35	0,010384615
4	-0,382	30	31,61	-1,61	0,053666667
5	-0,382	30	31,61	-1,61	0,053666667
6	-0,312	95	96,32	-1,32	0,013894737
7	0,062	75	74,74	0,26	0,003466667
8	-0,235	160	160,99	-0,99	0,0061875
9	-0,382	30	31,61	-1,61	0,053666667
10	-0,39	30	31,64	-1,64	0,054666667
11	0,883	3	-0,72	3,72	1,24
12	0,257	65	63,92	1,08	0,016615385
13	-0,484	105	107,04	-2,04	0,019428571
14	0,319	130	128,65	1,35	0,010384615
15	-0,562	40	42,37	-2,37	0,05925
16	0,631	45	42,34	2,66	0,059111111
17	-0,203	20	20,86	-0,86	0,043
18	-3,12	40	53,16	-13,16	0,329
19	0,07	75	74,71	0,29	0,003866667
20	-0,562	40	42,37	-2,37	0,05925
21	0,397	195	193,33	1,67	0,008564103
22	-0,757	50	53,19	-3,19	0,0638
23	0,639	250	247,3	2,7	0,0108
24	0,46	260	258,06	1,94	0,007461538
25	-0,094	290	290,4	-0,4	0,00137931
26	-0,094	290	290,4	-0,4	0,00137931
27	-0,172	225	225,73	-0,73	0,003244444
28	0,07	75	74,71	0,29	0,003866667
29	0,062	75	74,74	0,26	0,003466667
30	0,883	3	-0,72	3,72	1,24
31	3,166	45	31,64	13,36	0,296888889
32	0,883	3	-0,72	3,72	1,24
33	0,883	3	-0,72	3,72	1,24
34	-1,116	70	74,71	-4,71	0,067285714
35	-0,382	30	31,61	-1,61	0,053666667
36	-0,39	30	31,64	-1,64	0,054666667
37	-0,305	95	96,29	-1,29	0,013578947
38	-1,046	135	139,41	-4,41	0,032666667
39	2,401	225	214,87	10,13	0,045022222
40	0,883	3	-0,72	3,72	1,24
41	-5,498	30	53,19	-23,19	0,773
42	1,06	90	85,53	4,47	0,049666667

Cont...APPENDIX B

Diagnostic per case(a)					
Number of case	**Tip. Residual**	**Time**	**Predicted value**	**Residual**	**MRE**
43	0,132	140	139,44	0,56	0,004
44	3,189	250	236,55	13,45	0,0538
45	-0,998	200	204,21	-4,21	0,02105
46	-0,375	30	31,58	-1,58	0,052666667
47	-0,499	105	107,11	-2,11	0,020095238
48	-0,577	40	42,43	-2,43	0,06075
49	-0,57	40	42,4	-2,4	0,06
50	-0,382	30	31,61	-1,61	0,053666667
51	-0,195	20	20,82	-0,82	0,041
52	-1,046	135	139,41	-4,41	0,032666667
53	-0,055	150	150,23	-0,23	0,001533333
54	-0,04	150	150,17	-0,17	0,001133333
55	0,249	65	63,95	1,05	0,016153846
56	-0,57	40	42,4	-2,4	0,06
57	-0,008	10	10,03	-3,42E-02	0,003
58	-0,203	20	20,86	-0,86	0,043
59	-0,382	30	31,61	-1,61	0,053666667
60	-0,39	30	31,64	-1,64	0,054666667
61	-0,57	40	42,4	-2,4	0,06
62	-0,749	50	53,16	-3,16	0,0632
63	0,257	65	63,92	1,08	0,016615385
64	0,883	3	-0,72	3,72	1,24
65	0,07	75	74,71	0,29	0,003866667
66	-0,39	30	31,64	-1,64	0,054666667
67	-0,118	85	85,5	-0,5	0,005882353
68	0,03	215	214,87	0,13	6,04651E-04
69	-0,203	20	20,86	-0,86	0,043
70	-0,195	20	20,82	-0,82	0,041
71	-0,195	20	20,82	-0,82	0,041
72	0,436	55	53,16	1,84	0,033454545
73	-2,566	10	20,82	-10,82	1,082
74	-0,57	40	42,4	-2,4	0,06
75	-0,016	10	10,07	-6,60E-02	0,007
76	-0,749	50	53,16	-3,16	0,0632
77	0,883	3	-0,72	3,72	1,24
78	0,623	45	42,37	2,63	0,058444444
79	0,319	130	128,65	1,35	0,010384615
80	-0,562	40	42,37	-2,37	0,05925
81	-0,804	190	193,39	-3,39	0,017842105
82	-0,172	225	225,73	-0,73	0,003244444
83	-0,601	180	182,54	-2,54	0,014111111
84	-0,422	170	171,78	-1,78	0,010470588

Cont...APPENDIX B

Diagnostic per case(a)					
Number of case	**Tip. Residual**	**Time**	**Predicted value**	**Residual**	**MRE**
85	0,771	175	171,75	3,25	0,018571429
86	-0,203	20	20,86	-0,86	0,043
87	-0,297	95	96,25	-1,25	0,013157895
88	-0,936	60	63,95	-3,95	0,065833333
89	-0,57	40	42,4	-2,4	0,06
90	0,409	1	-0,72	1,72	1,72
91	-0,203	20	20,86	-0,86	0,043
92	0,99	25	20,82	4,18	0,1672
93	0,883	3	-0,72	3,72	1,24
94	-1,388	15	20,86	-5,86	0,390666667
95	-0,032	150	150,14	-0,14	9,33333E-04
96	0,409	1	-0,72	1,72	1,72
97	1,068	90	85,5	4,5	0,05
98	0,07	75	74,71	0,29	0,003866667
99	-0,188	20	20,79	-0,79	0,0395
100	3,915	5	-11,51	16,51	3,302
101	-0,008	10	10,03	-3,42E-02	0,003
102	1,083	90	85,43	4,57	0,050777778
103	-0,203	20	20,86	-0,86	0,043
104	0,883	3	-0,72	3,72	1,24
105	-0,016	10	10,07	-6,60E-02	0,007
106	0,883	3	-0,72	3,72	1,24
107	-0,016	10	10,07	-6,60E-02	0,007
108	0,883	3	-0,72	3,72	1,24
109	2,815	65	53,13	11,87	0,182615385
110	0,811	35	31,58	3,42	0,097714286
111	-0,203	20	20,86	-0,86	0,043
112	0,89	3	-0,76	3,76	1,253333333
113	-0,936	60	63,95	-3,95	0,065833333
114	1,13	155	150,23	4,77	0,030774194
115	-0,749	50	53,16	-3,16	0,0632
116	-0,094	290	290,4	-0,4	0,00137931
117	0,826	240	236,51	3,49	0,014541667
118	0,217	205	204,08	0,92	0,004487805
119	-0,601	180	182,54	-2,54	0,014111111
120	0,14	140	139,41	0,59	0,004214286
121	-0,749	50	53,16	-3,16	0,0632
122	-0,788	190	193,33	-3,33	0,017526316
123	-0,39	30	31,64	-1,64	0,054666667
124	-0,016	10	10,07	-6,60E-02	0,007
125	-0,133	85	85,56	-0,56	0,006588235
126	-0,195	20	20,82	-0,82	0,041

Cont...APPENDIX B

Diagnostic per case(a)					
Number of case	**Tip. Residual**	**Time**	**Predicted value**	**Residual**	**MRE**
127	-0,203	20	20,86	-0,86	0,043
128	-0,382	30	31,61	-1,61	0,053666667
129	0,883	3	-0,72	3,72	1,24
130	-0,016	10	10,07	-6,60E-02	0,007
131	-0,016	10	10,07	-6,60E-02	0,007
132	-0,203	20	20,86	-0,86	0,043
133	-0,203	20	20,86	-0,86	0,043
134	-0,195	20	20,82	-0,82	0,041
135	0,883	3	-0,72	3,72	1,24
136	0,312	130	128,69	1,31	0,010076923
137	-1,388	15	20,86	-5,86	0,390666667
138	-0,203	20	20,86	-0,86	0,043
139	-0,344	235	236,45	-1,45	0,006170213
140	0,015	215	214,94	6,38E-02	2,79070E-04
141	-0,22	160	160,93	-0,93	0,0058125
142	-0,016	10	10,07	-6,60E-02	0,007
143	0,883	3	-0,72	3,72	1,24
	MMRE=	0,236870			