FINDING RELATIONSHIPS BETWEEN EFFORT AND OTHER VARIABLES IN THE SEL

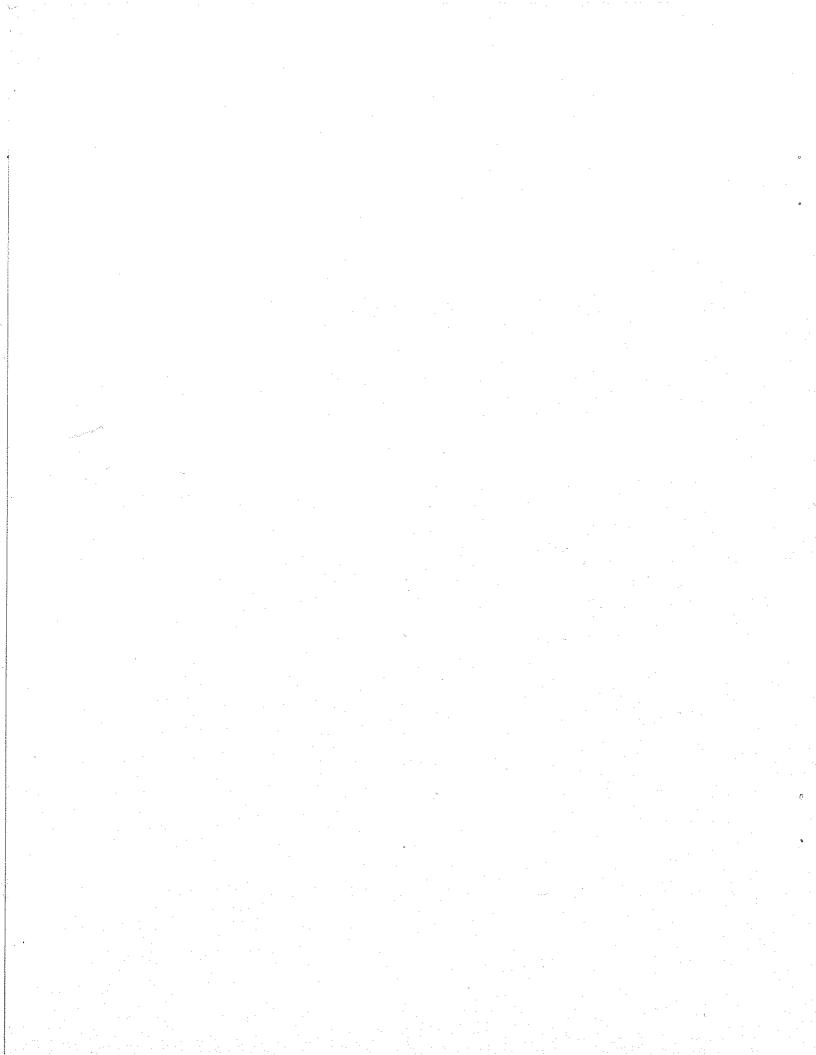
Victor R. Basili N. Monina Panlilio-Yap

Department of Computer Science University of Maryland College Park MD 20742

ABSTRACT

Estimating the amount of effort required for a software development project is one of the major aspects of resource estimation for that project. In this study, we examined the relationship between effort and other variables for 23 Software Engineering Laboratory projects that were developed for NASA/Goddard Space Flight Center. These variables fell into two categories: those which can be determined in the early stages of project development and may therefore be useful in a baseline equation for predicting effort in future projects, and those which can be used mainly to characterize or evaluate effort requirements and thus enhance our understanding of the software development process in this environment. The results of our analyses are presented in this paper.

^{*} Research for this study was supported in part by the National Aeronautics and Space Administration under grant NSG-5123 to the University of Maryland. Computer support was provided in part by the Computer Science Center at the University of Maryland.



1. INTRODUCTION

The estimation of resources required in the development of a software project is an issue of importance to managers. The development of useful models and equations for predicting the cost of a project is one of the major goals of software engineering. One of the ways of measuring cost is to measure the amount of effort and resources required for a project.

Several studies on measures of effort have been made and two basic approaches have been taken in these studies. [Wolverton 74], [Putnam 78], and [Boehm 81], among others, have developed generalized models which are then parameterized for a given environment in order to predict effort. The models are based upon data from at least one environment which is hoped to be typical or representative. [Walston & Felix 77], [Jeffery and Lawrence 79], [Basili & Freburger 81], [Boydston 84], etc., have collected data from several projects in a given environment and used these data to build models for characterizing or predicting effort in that environment, as we have in the Software Engineering Laboratory. Because of the differences in the environments, the types of projects and the data collected, [Bailey and Basili 81] have suggested that even generalized models are not necessarily transportable to other environments where a different set of factors come into play in different degrees.

[Bailey & Basili 81] have proposed a method for generating a resource estimation model for a particular organization based on data collected in that environment. These data would capture environmental factors and differences among projects which may have some impact on the software development process. The basic approach is as follows: A background equation is computed. The factors that could possibly explain the difference between the actual effort and the effort predicted by the background equation for the available data are analyzed. The model is then used to predict the effort required for a new project. The approach requires a local data base. If such a data base is not available then clearly one of the generalized models is best.

It has been suggested by [Boehm 84] that lines of code is not necessarily the best predictor for effort. In a study conducted at the IBM Santa Teresa Laboratory, [Boydston 84] has found that the number of modified modules is very strong statistically and is superior to lines of code as a single variable determinator of effort. Thus we are searching our summary statistics file to see if there are other variables that might be better to use, especially in a baseline equation.

This paper presents the results of some exploratory analysis on data collected in the Software Engineering Laboratory, a joint effort of NASA/Goddard Space Flight Center, Computer Sciences Corporation and the University of Maryland, which seeks to characterize and evaluate various models, metrics and software engineering practices to improve our understanding and management of both the software development process and the product. An attempt is made to find a model for effort as a function of various variables. This study also reexamines some of the relationships derived in an earlier study in the SEL by [Basili & Freburger 81] based on fewer data points.

2. BACKGROUND

Data were collected in the Software Engineering Laboratory from ground support software projects at NASA Goddard Space Flight Center. These projects were designed for similar applications. The code is written mostly in FORTRAN except for a small percentage written in macro assembler. Three sets of data were used in this study. One set (DS1) contained 23 data points. The other two sets were subsets of this. One of them (DS2) contained projects under 50 K lines of code. It had 15 data points. The other (DS3) contained projects consisting of 50 K or more lines of code. It had eight data points. One of the original projects included in DS3 was eliminated because it involved an unusually large amount of reused code. It was replaced by a large project which actually consisted of eight of the smaller projects in DS1. Appendix 2 shows the data used in each set.

Effort in this study is expressed in terms of staff-months. It consists of total programmer and management time for a given project. One staff-month of effort is defined as 160 staff-hours. Equations were derived with effort as the response variable. A list of the acronyms used is presented in Appendix 1. Table 1 shows the list of independent variables considered for regression. Definitions of terms used in the SEL may be found in [SEL-82-105] and the most important ones follow:

- 1. The total number of lines of code is the total number of lines of source code generated as a deliverable item for a project. It includes all executable, nonexecutable, and comment statements, whether newly coded or obtained from existing programs and library routines.
- 2. The number of new lines of code is the total number of lines of source code written by programmers for a given task. It excludes code taken from previously existing programs, but it includes comments, executable and non-executable statements.
- 3. The number of *modified lines* of code is the number of lines of previously developed code that has been changed for reuse in a new system.
- 4. The number of developed lines of code is defined in [Bailey & Basili 81] as the number of new lines of code plus twenty percent of the number of reused lines of code. System integration and full system test are accounted for by the 20% overhead.
- 5. The total number of *modules* in a project is the number of independently compilable units such as FORTRAN functions, subroutines and BLOCK DATA, or separately identifiable and retrievable components from an on-line library.
- 6. The number of new modules is the number of modules that are not reused from some previous project.
- 7. The number of *modified modules* is the number of previously developed modules that has been modified in some way for reuse in a new system.
- 8. A component is a named piece of a system. Examples are a separately compilable function, a functional subsystem, or a shared section of data such as a COMMON block.
- 9. The number of *computer runs* is determined by the computer accounting systems and includes every job submittal for batch systems and every terminal sign-on for interactive systems.
- 10. The number of pages of documentation consists of written material, excluding source code statements and comments embedded therein, that describes a

system or any of its components. It includes the program design document, development plan, test plans, system description, module descriptions and user's guide.

Newmodsq $(newmods^2)$ and newratio (newlines / newmods) were variables suggested by [Boydston 84]. They were found to have some significance on effort in studies made at the IBM Santa Teresa Laboratory.

The stepwise regression procedure of the Statistical Analysis System (SAS) package was initially used. More specifically, the maximum R^2 improvement (MAXR) technique was used for exploratory analysis. From the results, further analysis was done using the general linear models (GLM) procedure, and some plots were generated for single independent variable models that would possibly be useful in predicting effort during the early stages of development or models that showed a strong correlation between effort and the independent variable whose value cannot be determined early in the development. The latter cannot be used for resource estimation purposes but may be useful in characterizing effort in the environment.

The maximum R^2 improvement technique is considered almost as good as all possible regressions. [SAS 82]. It tries to find the best one-variable model, the best two-variable model, and so on. Initially, it finds the one-variable model that yields the highest value for R^2 . Another variable which gives the greatest increase in R^2 is added. After obtaining the two-variable model, each of the variables in the model is compared to each variable that is not in the model. For each comparison, the technique decides if deleting a variable and replacing it with the other variable results in an increase in R^2 . When all possible switches have been compared, the one producing the maximum increase in R^2 is made. At this point, comparisons are made again. This continues until the technique can no longer find any switch that could increase R^2 . Therefore the two-variable model generated is considered the best that the maximum R^2 improvement technique can find. One more variable is then added to the model, and the comparing-

Table 1 - Variables Considered for Determining Effort

number of developed lines of code
number of pages of documentation
number of modified lines of code
number of modified modules
number of new lines of code
number of new modules
ratio of new lines of code to new modules
number of source code changes (versions)
number of components
number of computer runs
total number of lines of code
total number of modules

and-switching process is repeated until the best three-variable model is found. When there are no more variables that can be added to the model to increase the value of R^2 , the procedure stops.

3. ANALYSES AND RESULTS

Each data set was given several sets of candidate independent variables to be used in generating a model. The response variable used was effort.

For each set of candidate independent variables, the following steps were taken.

- 1. Run STEPWISE/MAXR to generate the best n-variable model, n = 1, 2, ...
- 2. Leave out of further consideration those models with significance probability (Prob>F)>0.05. Only consider those models where (Prob>F)<=0.05 for the entire model and for each of the independent variables included in the model.
- 3. Disregard models where the ratio of the number of data points to the number of independent variables is less than 5 for DS1 and DS2. For DS3, because of the limited number of data points, consider models with up to 3 independent variables.
- 4. Disregard models with $R^2 < 0.5$. Preferably, R^2 should be > = 0.7 so that the model accounts for at least 70% of variation of effort in the model.
- 5. Disregard models with n variables where the increment of R^2 over that of the model with n-1 variables is very small. Do this for i = 2,...,n.

Across sets of candidate independent variables

- 1. Avoid models with higher-ordered terms.
- 2. Select model from the set with the greatest number of candidate independent variables originally supplied.
- 3. Select the model with the highest value of R^2 for the smallest number of variables.

The general linear models (GLM) procedure of SAS was used to examine in closer detail some of the more interesting one-variable models for each data set. Overlaid plots of predicted and actual values of effort versus the independent variables were generated. Plots of the residuals versus the independent variables were also produced.

3.1. All Projects

The first data set consists of 23 projects ranging in size from 2.1 KLOC to 111.9 KLOC. The mean is 33.3 KLOC and the standard deviation is 32.5 KLOC. The number of modules ranges from 23 to 535 with a mean value of 198 and a standard

deviation of 172. Effort for these projects ranges from 2.4 to 121.7 staff months. Mean effort is 40.9 staff-months and the standard deviation is 40.4 staff-months. Because the ranges are wide and the standard deviations are large, we subsequently formed the two smaller data sets and analyzed them separately.

Of all the sets of candidate independent variables used to generate a model of effort for this data set, only those which gave reasonable and interesting results are presented here. Initially, the set of candidate independent variables consisted only of newlines, newmods, modlines, and modmods. These are analogous to variables used by [Boydston 84]. They can be determined in the early stages of project development and may therefore have predictive value. The one-variable equation that resulted is

$$Effort = 5.497 + 1.500 newlines \tag{1}$$

$$R^2 = 0.795$$
 $F = 81.65$ $Prob > F = 0.0001$

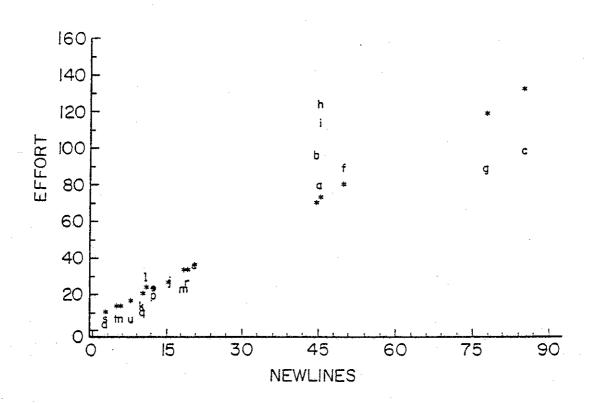
The standard error of estimate (SEE) for the slope of this equation is 0.166. Adding newratio to the set of candidate independent variables yielded the same result. Figure 1 shows a plot of actual effort versus newlines for the different projects. The letters in the figure represent the different projects. Some observations are hidden due to overlap in values. The figure also shows the corresponding points predicted by equation (1). These are represented by asterisks (*) in the plot. Figure 2 is a plot of the residuals versus newlines for this equation. As in Figure 1, the letters represent the different projects in the data set.

These plots show a few points for which there is a large discrepancy between the actual and the predicted values of effort. Projects h and i for which the equation underpredicts the value of effort were both developed when there was a major change in the environment. A more reliable machine and more computer terminals were installed. There was quicker turnaround. The staff were turned loose on the computer. However, the level of experience of the staff for both these projects was lower than for most of the other projects in the study. Project h was a problem project. It did not have enough experienced staff to begin with and staffing adjustments had to be made in midstream. It was very late compared to other projects and was undertested. Project i was also a potential problem project, but was given more attention because of the unhappy experience with project h and was fortunately straightened out sooner. Projects c and g for which the equation (1) overpredicts effort were both developed with more experienced staff.

Because the number of developed lines was originally found by [Bailey & Basili 81] to be the best predictor of effort in their meta-model for the SEL, it was added to the set of candidate independent variables. The resulting equation is

$$Effort = 4.372 + 1.430 \ devlines$$
 (2)
 $R^2 = 0.808 \ F = 88.30 \ Prob > F = 0.0001$

The SEE for the slope of this equation is 0.152. Figures 3 and 4 show the plot of actual and predicted values of effort versus developed lines and the plot of residuals versus



- 1. Actual effort vs. newlines---symbol used is letter code of project.
- 2. Predicted effort vs. newlines---symbol used is *.
- 3. 9 observations hidden.

Figure 1 - Effort vs. newlines for DS1.

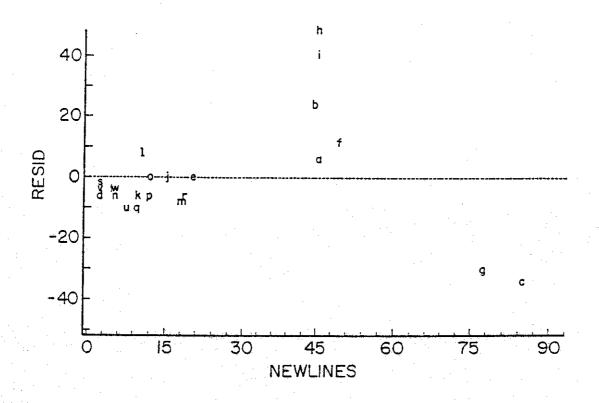


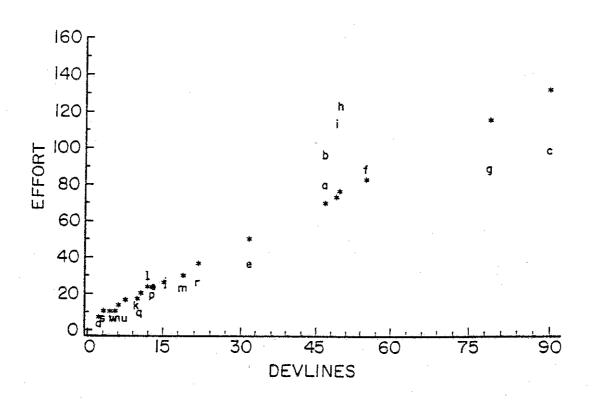
Figure 2 - Effort residuals vs. newlines for DS1.

developed lines respectively for equation (2). The outlier points are the same as those for equation (1). There is little difference in these two models and they indicate that the number of developed lines is at least as good as the number of new lines for predicting effort in the SEL.

Another set of models was generated using newlines, newmods, modlines, modmods and the squares of each of these. [Boydston 84] found that there is a quantified square root trade-off between the number of new modules and the amount of new code per module and thus included a newmods² term in one of his effort equations. We sought to discover whether or not the inclusion of the squared terms would have any effect on the effort model. The one-variable model that resulted is the same as (1) above. Two and three variable models were also generated as follows:

Effort = -11.938 + 3.427 newlines - 0.025 newlinsq (3)

$$R^2 = 0.916 \ F = 109.70 \ Prob > F = 0.0001$$



- 1. Actual effort vs. devlines---symbol used is letter code of project.
- 2. Predicted effort vs. devlines---symbol used is *.
- 3. 4 observations hidden.

Figure 3 - Effort vs. devlines for DS1.

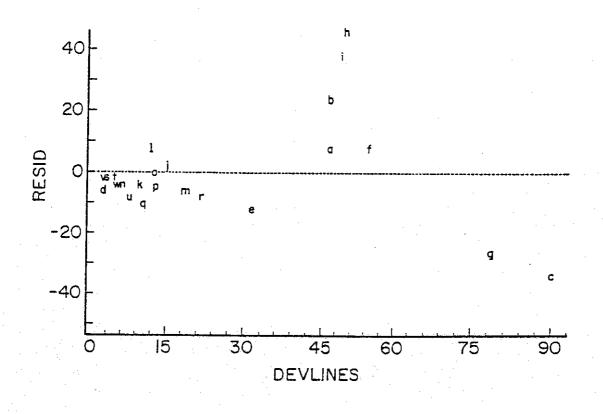


Figure 4 - Effort residuals vs. devlines for DS1.

Effort = -13.740 + 3.258 newlines + 0.355 mod mods

$$-0.028$$
 newlinsq

$$R^{2} = 0.933 F = 88.53 Prob > F = 0.0001$$
(4)

There is a substantial increase in R^2 in going from equation (1) to (3), but not from (3) to (4). This suggests that the quadratic equation (3) may be a much better model than (1), but the inclusion of the additional term modmods does not improve the model tremendously and only adds to the complexity.

Similarly adding the square of developed lines to equation (2) to parallel equation (3) yields the following model:

Effort = -10.588 + 2.992 devlines - 0.020 devlinsq (5)

$$R^2 = 0.900 \ F = 89.81 \ Prob > F = 0.0001$$

This result shows a considerable improvement in R^2 over equation (2). Comparing equations (3) and (5) again shows little difference in the predictive power of new lines and developed lines.

In this study, in addition to predictive models of effort, we also sought relationships between effort and other variables in our database. Models for characterizing and evaluating effort could enhance our understanding of the software development process in our environment. A set of models using as candidate independent variables all the variables in Table 1 with the exception of newratio was generated. Newratio was already found to be insignificant so it was excluded. Models with up to three variables were reasonable in this case and are presented here.

$$Effort = 9.951 + 0.008 \ numruns$$

$$R^{2} = 0.895 \ F = 179.30 \ Prob > F = 0.0001$$
(6)

$$Effort = 3.384 + 0.104 \ newmods + 0.006 \ numruns$$
 (7)

$$R^2 = 0.939$$
 $F = 154.57$ $Prob > F = 0.0001$

$$Effort = 4.484 - 0.637 \ mod mod s + 0.963 \ devlines$$

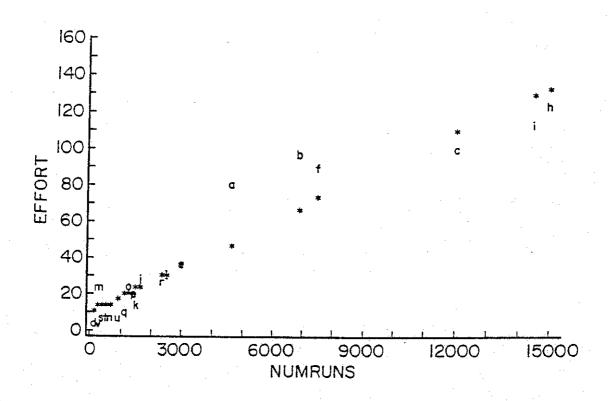
$$+ 0.007 \ numruns$$

$$(8)$$

$$R^2 = 0.978 \ F = 285.34 \ Prob > F = 0.0001$$

The one-variable model shows a very strong relationship between the number of runs in the project and the amount of effort. Numruns, by itself, accounts for 90% of the variation in effort. The SEE for the slope of equation (6) is 0.0006. Figures 5 and 6 show the plots of actual and predicted effort versus number of runs and residuals versus number of runs respectively for equation (6). For project a there is more actual effort per number of runs. It is one of the earliest projects included in this study. The developers were relatively inexperienced. This project was also characterized by staffing and management problems early in the project and serious staffing changes. On the other hand, project b was developed by experienced staff. In this case, the actual effort is also higher than that predicted by the equation, probably because the developers were more thorough and purposely put in more effort per run.

To see what other variables correlate well with effort, the number of runs was excluded from the set of candidate independent variables. The following one-variable and four-variable models were generated:



- 1. Actual effort vs. numruns---symbol used is letter code of project.
- 2. Predicted effort vs. numruns---symbol used is *.
- 3. 5 observations hidden.

Figure 5 - Effort vs. numruns for DS1.

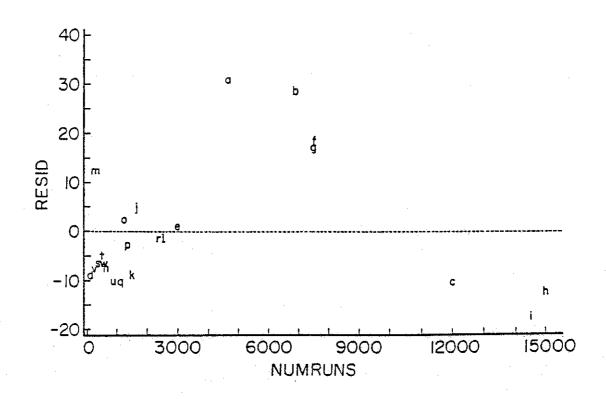


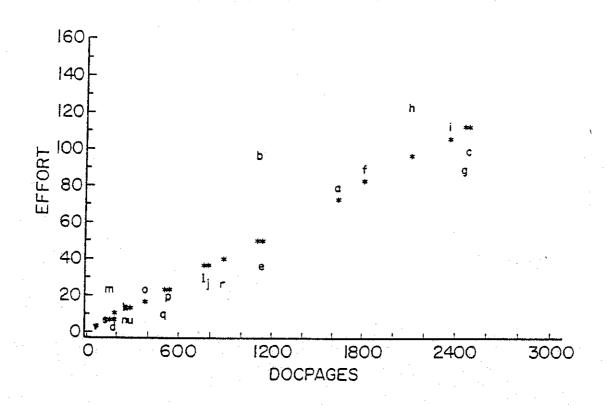
Figure 6 - Effort residuals vs. numruns for DS1.

Effort = 0.581 + 0.045 docpages (9)

$$R^2 = 0.871 \ F = 141.82 \ Prob > F = 0.0001$$

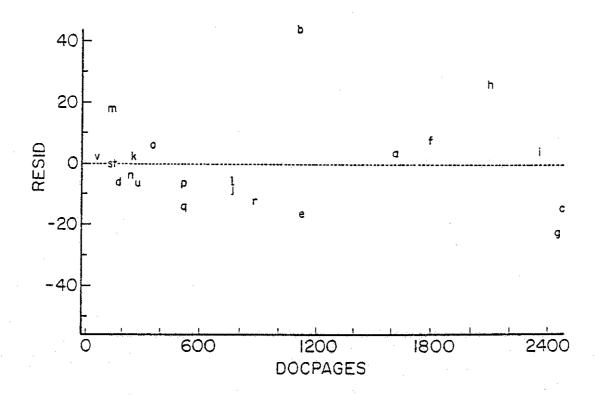
Effort = 2.634 - 0.346 newmods - 5.076 modlines
+ 0.045 docpages + 0.066 numchngs
$$R^{2} = 0.930 F = 59.99 Prob > F = 0.0001$$
(10)

The one-variable equation shows the strong relationship that characterizes effort across these projects and pages of documentation. Like number of runs, however, this relationship cannot be used for predictive purposes since the number of pages of documentation cannot really be determined early in the project. The SEE for the slope of equation (9) is 0.0038. Figures 7 and 8 show plots of effort and residuals versus pages of documentation respectively for equation (9).



- 1. Actual effort vs. docpages---symbol used is letter code of project.
- 2. Predicted effort vs. docpages---symbol used is *.
- 3. 3 observations hidden.

Figure 7 - Effort vs. docpages for DS1.



Note: 1 observation hidden.

Figure 8 - Effort residuals vs. docpages for DS1.

Because of the high correlation of both number of runs and pages of documentation with effort, we investigated whether or not a predictor variable like the number of developed lines could be used to predict their values. If we could determine ahead of time the number of runs a project would entail, we could use this information to allocate computer time. Similarly, if we could obtain a good estimate of the number of pages of documentation during the early stages of project development, we could get the publications group ready. We obtained the following result for the number of runs:

Numruns = -108.274 + 150.879 devlines (11)

$$R^2 = 0.686 \ F = 45.90 \ Prob > F = 0.0001$$

The SEE for the slope of equation (11) is 22.269. We can see from this that the number of developed lines does not explain more than 69% of the variation in number of runs. We obtained a much better result for pages of documentation:

Docpages = 99.143 + 30.895 devlines (12)

$$R^2 = 0.892 \ F = 173.24 \ Prob > F = 0.0001$$

The SEE for the slope of equation (12) is 2.347. Figure 9 shows a plot of actual and predicted pages of documentation versus number of developed lines based on equation (12). Figure 10 shows the residuals plotted against developed lines. The number of developed lines accounts for almost 90% of the variation in pages of documentation. [Basili & Freburger 81] obtained the following equation from a smaller set of projects in the SEL:

$$Doc = 34.7 (DL^{0.93)}$$

where

Doc = pages of documentation

DL = number of developed lines

They noted that the relationship is approximately linear and our result tends to support this observation.

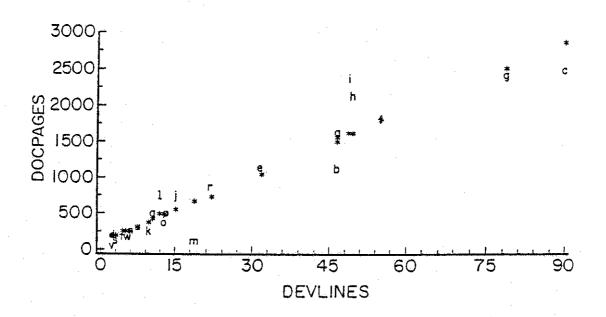
Deleting both numruns and docpages from the independent variable set yields two reasonable models for effort. The one-variable model is the same as equation (2) above. The four-variable model generated is

Effort =
$$5.433 - 1.082$$
 newmods + 22.376 newlines
+ 0.854 totmods - 20.476 devlines

$$R^{2} = 0.926 F = 56.24 Prob > F = 0.0001$$
(13)

Many other combinations of variables were used but they either failed to produce any interesting results or they yielded the same results as the case above which included all variables in Table I with the exception of newratio. The number of runs is the independent variable most highly correlated with effort when all 23 projects are considered. It may be an excellent measure of the complexity of the project, the quality of the development, the quality of the product, the amount of testing involved, the level of structure or disorganization of project management, and a variety of other factors.

3.2. Projects Under 50 K Lines of Code



- 1. Actual docpages vs. devlines---symbol used is letter code of project.
- 2. Predicted docpages vs. devlines---symbol used is *.
- 3. 2 observations hidden.

Figure 9 - Docpages vs. devlines for DS1.

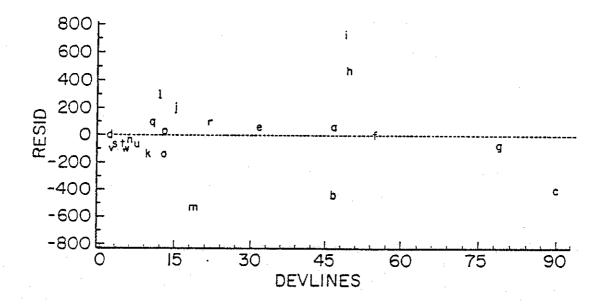


Figure 10 - Docpages residuals vs. devlines for DS1.

There are 15 projects with less than 50 K total lines of code. They range in size from 2.1 to 32.8 KLOC with a mean of 11.9 KLOC and a standard deviation of 8.1 KLOC. There are from 23 to 263 modules. The mean number is 91 and the standard deviation is 63. Effort ranges from 2.4 to 29.0 staff-months with a mean of 14.6 and a standard deviation of 9.5.

In all cases where the number of new lines was included in the set of candidate independent variables, the following model was generated:

Effort = 0.877 + 1.535 newlines (14)

$$R^2 = 0.802 \ F = 52.71 \ Prob > F = 0.0001$$

This is so even where the number of runs was included. This equation is selected by the STEPWISE/MAXR technique as the best single-variable effort equation for the smaller projects. It has predictive power since the number of new lines can be estimated early in the development of the project. The SEE for equation (14) is 0.211. The plots of effort versus new lines and residuals versus newlines for this equation are shown on Figures 11

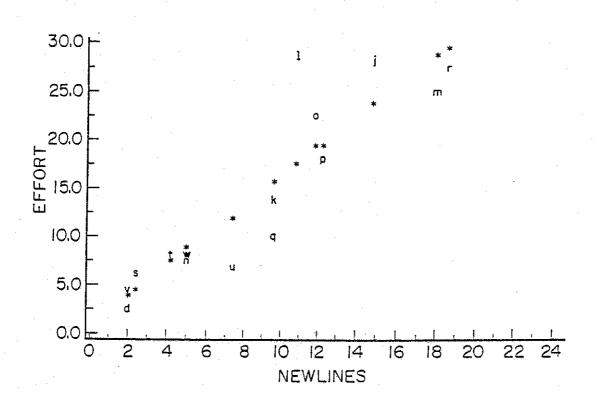
and 12 respectively.

Where newlines was not included in the set of candidate independent variables, the number of developed lines was selected by the technique. The equation generated is

Effort = 1.013 + 1.423 devlines (15)

$$R^2 = 0.797 \ F = 50.92 \ Prob > F = 0.0001$$

The SEE for the slope of this equation is 0.199. Figures 13 and 14 respectively show plots of effort and residuals versus the number of developed lines. They are very similar to Figures 11 and 12. In the absence of new lines or developed lines, the number of total lines was selected as the predictor variable.



- 1. Actual effort vs. newlines---symbol used is letter code of project.
- 2. Predicted effort vs. newlines---symbol used is *.
- 3. 2 observations hidden.

Figure 11 - Effort vs. newlines for DS2.

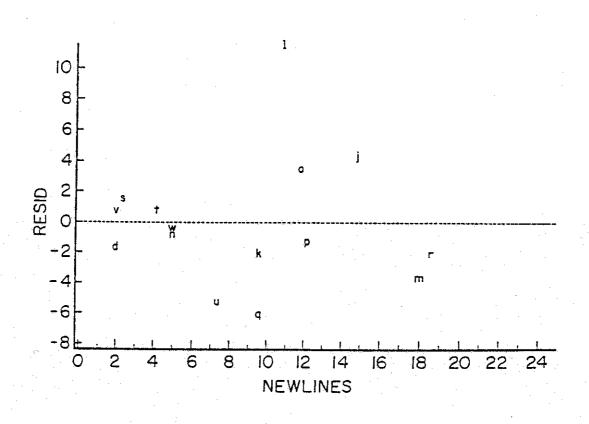
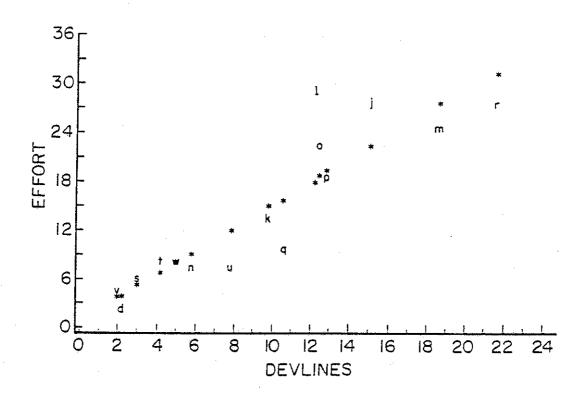


Figure 12 - Effort residuals vs. newlines for DS2.



- 1. Actual effort vs. devlines---symbol used is letter code of project.
- 2. Predicted effort vs. devlines---symbol used is *.

Figure 13 - Effort vs. devlines for DS2.

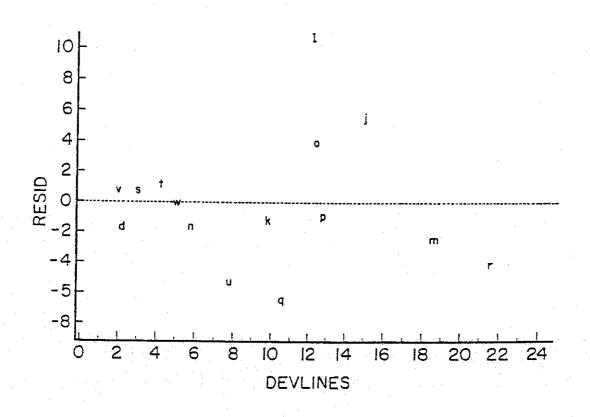


Figure 14 - Effort residuals vs. devlines for DS2.

The two-variable equation generated in most attempted cases is

Effort = -1.185 + 0.108 newmods + 0.009 numruns (16)

$$R^2 = 0.890 \ F = 48.53 \ Prob > F = 0.0001$$

The number of runs entered the model and the number of new modules replaced the number of new lines. None of the other models generated were reasonable.

3.3. Projects with 50 K Lines of Code or More

There are 8 projects in this data set. The smallest consists of 50.9 KLOC and the largest is 111.9 KLOC. The mean size is 75.2 KLOC with a standard deviation of 19.9 KLOC. The number of modules ranges from 201 to 604 with a mean value of 427 and a

standard deviation of 139. Effort ranges from 78.7 to 121.7 staff-months with a mean of 97.7 and a standard deviation of 13.5.

In all cases where the number of runs was included in the set of candidate independent variables, the following was generated as the best one-variable model:

$$Effort = 66.868 + 0.003 \ numruns$$
 (17)
 $R^2 = 0.878 \ F = 43.27 \ Prob > F = 0.0006$

The standard error of estimate for the slope of this equation is 0.0005. Figures 15 and 16 respectively show plots of effort and residuals versus the number of runs for the larger projects.

All models which excluded number of runs from the set of independent variables yielded the following as the only reasonably good equation:

$$Effort = 122.220 + 1.088 \ mod mod s - 0.960 \ newlines$$

$$-3.883 \ mod lines$$
(18)

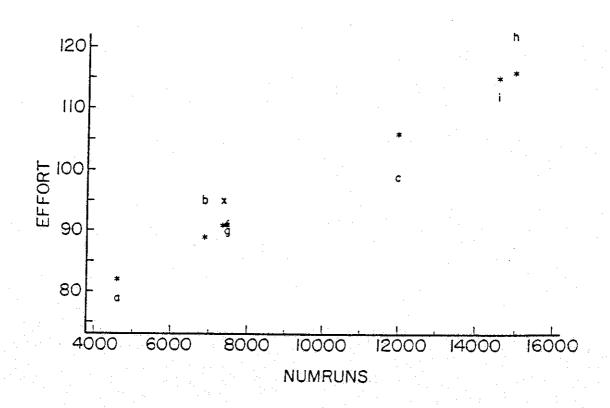
$$R^2 = 0.917$$
 $F = 14.78$ $Prob > F = 0.0125$

For the larger projects, the number of modified modules is always the first independent variable selected for entry into the model. It seems to be a better predictor of effort for this environment than lines of code, whether new, developed or total, but not much better. It explains only 17% of the variation in effort. It seems that none of the variables which can be determined early in project development is a good single predictor of effort in larger projects.

There were no good two-variable models generated. Considering there are only 8 data points, we should exercise caution in using equation (18) for predictive purposes.

4. SUMMARY

As was shown in [Bailey & Basili 81], developed lines of code is a good overall predictor of effort across all the projects considered in this study. It is one of the variables that can be estimated early in the project development and can thus be used to predict the effort requirements. For projects under 50 KLOC, the number of new lines was found to be the most significant predictor of effort whenever it was included in the set of candidate independent variables. The number of developed lines similarly predicts effort well for the smaller projects. The number of modified modules was not found to be most significant as a single predictor of effort in the Software Engineering Laboratory data. This differs from the result obtained by [Boydston 84] in the IBM Santa Teresa Laboratory environment. Although the amount of code modification in the SEL was by no means small, it probably was not sufficient to show significance. For the projects that



- 1. Actual effort vs. numruns---symbol used is letter code of project.
- 2. Predicted effort vs. numruns---symbol used is *.
- 3. 1 observation hidden.

Figure 15 - Effort vs. numruns for DS3.

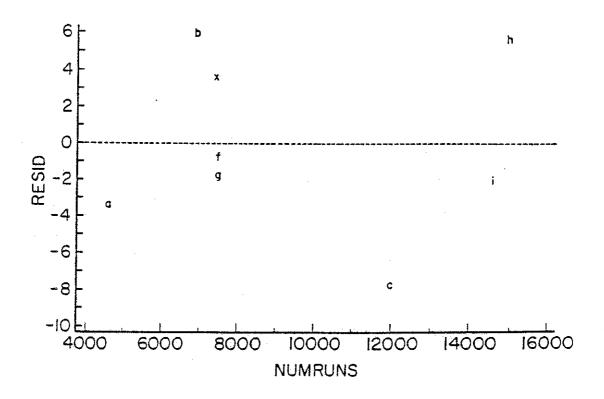


Figure 16 - Effort residuals vs. numruns for DS3.

contained 50 KLOC or more, the number of modified modules is a better single predictor of effort than any of the line measures. It produced a better model of effort than new-lines, devlines, modlines or totlines. However, it only accounts for 17% of the variation in effort and is clearly not by itself a good predictor of effort.

There is a high correlation between effort and number of runs overall and in the dataset containing projects that are at least 50 KLOC. There is also a high correlation between effort and pages of documentation across all the projects included in this study. Although these variables cannot be determined in the early stages of project development and therefore cannot be used for predictive purposes, they are nevertheless valuable for explaining and evaluating effort requirements in a project.

Only linear models, for the most part, were attempted for the sake of simplicity in this exploratory study. It would be premature at this point to select one of the above models as the best one to characterize, evaluate or predict effort. The selected model must be subjected to some test of stability. In this study, we were unable to find suitable substitutes for lines of code in a baseline equation to predict effort. However, this study does give us some indication of other variables in the SEL database which are

highly correlated with effort. From the non-predictive models generated, we can obtain valuable insight into the software development process in our environment.

5. RESEARCH DIRECTIONS

It may be worthwhile to investigate relationships that may exist among the other variables available in the SEL database. Such relationships may improve our understanding of the software product and its underlying development process. For example, the number of runs could be further examined to see if the data sets get partitioned into projects with high frequency of runs and projects with low frequency of runs (complex versus simple, under-managed versus well-managed, well-tested versus little-tested, etc.).

Nonlinear models could also be explored for these data sets. It is possible that the multi-variable models are curvilinear or ellipsoid rather than planar. The presence of negative intercepts may be due to a poor fit.

Other approaches to the basic meta-model [Bailey & Basili 81] could be investigated. Factor analysis could be used to isolate groups of environmental attributes that work together and to obtain good independent variables that could be used in multiple regression analysis. The principal components technique could also be used instead of factor analysis. If it is possible to justify the use of more than one independent variable to predict the value of the dependent variable in the background equation, greater accuracy may be obtained.

ACKNOWLEDGEMENTS

The authors are grateful to Mr. Frank McGarry of NASA/Goddard Space Flight Center and Dr. Jerry Page of Computer Sciences Corporation for their invaluable help in providing the data for this study. They also wish to thank the members of their research group at the University of Maryland and Dr. C. Frank Starmer of Duke University who have provided assistance of one form or another in the preparation of this paper.

REFERENCES

[Bailey & Basili 81]

Bailey, John W. and Victor R. Basili, "A Meta-Model for Software Development Resource Expenditures", Proceedings, Fifth International Conference on Software Engineering, 1981, pp. 107-116.

[Basili & Freburger 81]

Basili, V. R. and K. Freburger, "Programming Measurement and Estimation in the Software Engineering Laboratory", Journal of Systems and Software, Vol. 2, No. 1, 1981, pp. 47-57.

Boehm 81

Boehm, Barry W., Software Engineering Economics, Prentice-Hall, Englewood Cliffs, New Jersey, 1981.

[Boehm 84]

Boehm, Barry W., "Software Engineering Economics", IEEE Transactions on Software Engineering 10, No. 1, 1984.

[Boydston 84]

Boydston, Robert E., "Programming Cost Estimate: Is It Reasonable?", Proceedings of the Seventh ICSE, IEEE Computer Society Press, March 26-29 1984, pp. 153-159.

[Jeffery & Lawrence 79]

Jeffery, D. R. and M. J. Lawrence, "An Inter-organizational Comparison of Programming Productivity", Department of Information Systems, University of New South Wales, 1979.

[Putnam 78]

Putnam, L., "A General Empirical Solution to the Macro Software Sizing and Estimating Problem", IEEE Transactions on Software Engineering 4, No. 4, 1978.

[SAS 82]

SAS Institute Inc., SAS User's Guide: Statistics, 1982 Edition, Cary, NC: SAS Institute Inc., 1982.

[SEL-82-105]

"Glossary of Software Engineering Laboratory Terms", T. A. Babst, F. E. McGarry, and M. G. Rohleder, NASA/Goddard Space Flight Center, October 1983.

[SEL-83-001]

"An Approach to Software Cost Estimation", F. E. McGarry, G. Page, D. N. Card, et al., NASA/Goddard Space Flight Center, February 1984.

[Walston & Felix 77]

Walston, C. E. and C.P. Felix, "A Method of Programming Measurement and Estimation", IBM Systems Journal 16, No. 1, 1977.

[Wolverton 74]

Wolverton, R., "The Cost of Developing Large Scale Software", IEEE Transactions on Computers 23, No. 6, 1974.

Appendix 1 - List of Acronyms

devlines devlinsq devlines? docpages modlines modlines modmods modmodsq newlines newlines newlines newlines newlines newmodsq number of new modules newmodsq number of source code changes (versions) numcomps numruns projecde projname totlines total number of lines of code (KLOC) total number of modules allhrs mgmthrs proghrs + mgmthrs + servhrs management work time (tenths of an hour) servhrs mention newlines? number of modified modules modlines? number of new lines of code (KLOC) newlines? number of new modules number of source code changes (versions) number of components number of computer runs project code project code project code project name total number of lines of code (KLOC) total number of modules allhrs management work time (tenths of an hour) proghrs + mgmthrs proghrs programmer work time (tenths of an hour) servhrs	Acronym	Description
devlinsq docpages number of pages of documentation number of modified lines of code (KLOC) modlinsq modlines² number of modified modules modmodsq newlines number of new lines of code (KLOC) newlinsq number of new lines of code (KLOC) newlines² number of new modules newmodsq newmods² newmods² newmods² newmods² newlines / newmods number of source code changes (versions) numcomps number of components number of computer runs project code projname project code project code project name totlines total number of lines of code (KLOC) total number of modules allhrs mgmthrs proghrs programmer work time (tenths of an hour) proghrs		
docpages modlines modlines modlinsq modlinsq modmods modmodsq newlines newlines newmods newmodsq newratio numchngs numcomps numruns projcode projname totlines modmods allhrs mgmthrs proghrs proghrs number of pages of documentation number of modified lines of code (KLOC) modlines² number of modified modules modmods² number of new lines of code (KLOC) newlines² number of new modules newmods² newmods² newlines / newmods numcods number of components number of components number of computer runs project code project code project name total number of lines of code (KLOC) total number of modules proghrs + mgmthrs + servhrs management work time (tenths of an hour) proghrs programmer work time (tenths of an hour)	devlines	newlines $+0.2*$ (totlines - newlines) (KLOC)
modlines modlinsq modlines² modmods modmodsq newlines newlines newmods newmodsq newlines² newmods newmodsq newmodsq newmodsq newratio number of new lines of code (KLOC) newlines² newmods newmodsq newmods² newratio numchngs number of source code changes (versions) numcomps numcomps numruns projcode projname totlines total number of lines of code (KLOC) total number of modules allhrs mgmthrs mgmthrs proghrs number of modified lines of code (KLOC) number of modules number of new lines of code (KLOC) newlines² number of new modules newlines² number of source code changes (versions) number of components number of computer runs project code projname total number of lines of code (KLOC) total number of modules proghrs + mgmthrs + servhrs mgmthrs proghrs programmer work time (tenths of an hour)	devlinsq	devlines ²
modlinsq modmods modmodsq newlines newlines newmods newmodsq newratio numchngs numcomps numcomps numcode projname totlines total number of modified modules modmods² number of new lines of code (KLOC) newlines² number of new modules newmods² newmods² newlines / newmods numchngs number of source code changes (versions) numcomps number of components number of components project code projname total number of lines of code (KLOC) total number of modules allhrs mgmthrs mgmthrs management work time (tenths of an hour) proghrs programmer work time (tenths of an hour)	docpages	number of pages of documentation
modmods modmodsq newlines newlines newmods newmodsq newmodsq newmodsq newmodsq newmodsq newmodsq newmodsq newmodsq newmodsp numchngs numcomps numcomps numcomps numcomps numcomps numcomps projcode projname totlines total number of modified modules newlines of code (KLOC) newlines of new modules newmods of newmods numcomps number of source code changes (versions) number of components number of computer runs project code project code project name totlines total number of lines of code (KLOC) totmods allhrs mgmthrs mgmthrs mgmgmthrs mgnagement work time (tenths of an hour) proghrs programmer work time (tenths of an hour)	modlines	number of modified lines of code (KLOC)
modmodsq newlines newlinsq newmods newmods newmodsq newratio numchngs numcomps numcomps numruns projcode projname totlines total number of modules allhrs mgmthrs proghrs modmods² number of new lines of code (KLOC) newlines² number of new modules newmods² newmods² newmods² newmods² newmods² newmods² newmods	modlinsq	$mod lines^2$
newlines newlinsq newmods newmods newmodsq newmods² newmods² newlines / newmods numchngs numcomps numcomps numruns projcode projname totlines total number of lines of code (KLOC) totmods number of new modules newmods² newlines / newmods number of source code changes (versions) number of components number of computer runs project code projname total number of lines of code (KLOC) totmods allhrs proghrs + mgmthrs + servhrs mgmthrs proghrs programmer work time (tenths of an hour) proghrs programmer work time (tenths of an hour)	modmods	number of modified modules
newlinsq newmods newmodsq newmods² newmods² newmods² newlines² newmods² newmods² newlines / newmods numchngs number of source code changes (versions) numcomps number of components numruns projcode projname totlines total number of lines of code (KLOC) totmods allhrs mgmthrs mgmthrs mgmthrs proghrs programmer work time (tenths of an hour) programmer work time (tenths of an hour)	modmodsq	modmods ²
newmodsq newmodsq newmods² newmods² newmods² newmods number of newmods numchngs number of source code changes (versions) numcomps number of components number of computer runs project code projname totlines total number of lines of code (KLOC) total number of modules allhrs mgmthrs mgmthrs mgmghrs proghrs programmer work time (tenths of an hour) programmer work time (tenths of an hour)	newlines	number of new lines of code (KLOC)
newmodsq newmods² newmods² newmods² numchngs numcomps number of source code changes (versions) number of components number of computer runs project code projname totlines total number of lines of code (KLOC) totmods allhrs mgmthrs mgmthrs mgmthrs proghrs programmer work time (tenths of an hour) programmer work time (tenths of an hour)	newlinsq	newlines ²
newratio numchngs numcomps numruns projcode projname totlines allhrs mgmthrs mgmthrs proghrs newlines / newmods number of source code changes (versions) number of components number of computer runs project code project name total number of lines of code (KLOC) total number of modules proghrs + mgmthrs + servhrs management work time (tenths of an hour) proghrs programmer work time (tenths of an hour)	newmods	number of new modules
numchngs numcomps numruns projcode projname totlines totmods allhrs mgmthrs proghrs proghrs numchngs number of source code changes (versions) number of components number of computer runs project code project name total number of lines of code (KLOC) total number of modules proghrs + mgmthrs + servhrs management work time (tenths of an hour) proghrs programmer work time (tenths of an hour)	newmodsq	newmods ²
numcomps numruns projcode projname totlines total number of computer runs proghrs + mgmthrs proghrs programmer work time (tenths of an hour) proghrs number of components number of computer runs project code project name total number of lines of code (KLOC) total number of modules	newratio	newlines / newmods
numruns project code projname totlines total number of lines of code (KLOC) totmods allhrs mgmthrs proghrs proghrs programmer work time (tenths of an hour) proghrs programmer work time (tenths of an hour)	numchngs	number of source code changes (versions)
project code projname totlines total number of lines of code (KLOC) totmods allhrs mgmthrs prmghrs proghrs proghrs proger management work time (tenths of an hour) proghrs programmer work time (tenths of an hour)	numcomps	number of components
projname totlines total number of lines of code (KLOC) totmods total number of modules allhrs proghrs + mgmthrs + servhrs management work time (tenths of an hour) proghrs programmer work time (tenths of an hour)	numruns	number of computer runs
totlines total number of lines of code (KLOC) totmods total number of modules allhrs proghrs + mgmthrs + servhrs management work time (tenths of an hour) prmghrs proghrs programmer work time (tenths of an hour)	projcode	project code
totmods total number of modules allhrs proghrs + mgmthrs + servhrs mgmthrs management work time (tenths of an hour) prmghrs proghrs programmer work time (tenths of an hour)	projname	project name
allhrs proghrs + mgmthrs + servhrs management work time (tenths of an hour) proghrs programmer work time (tenths of an hour)	totlines	total number of lines of code (KLOC)
mgmthrs management work time (tenths of an hour) prmghrs proghrs programmer work time (tenths of an hour)	totmods	total number of modules
mgmthrs management work time (tenths of an hour) prmghrs proghrs + mgmthrs programmer work time (tenths of an hour)	allhrs	proghrs + mgmthrs + servhrs
prmghrs proghrs + mgmthrs programmer work time (tenths of an hour)	mgmthrs	
, , , , , , , , , , , , , , , , , , , ,	prmghrs	-
servhrs services work time (tenths of an hour)	proghrs	programmer work time (tenths of an hour)
	servhrs	services work time (tenths of an hour)
allefrt allhrs / 1600 (staff-months)	allefrt	allhrs / 1600 (staff-months)
effort prmghrs / 1600 (staff-months)	effort	, , , , , , , , , , , , , , , , , , ,

Appendix 2 - Data Used in the Analyses

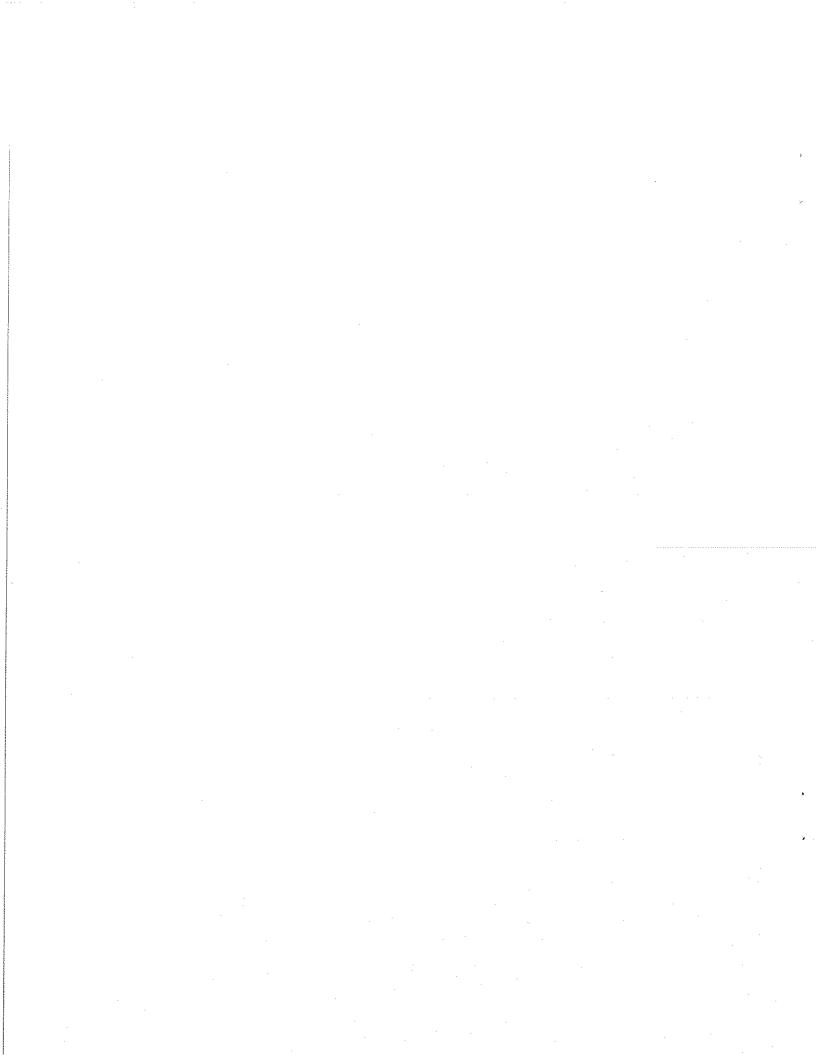
095	rs t t 3 j			N295005	1001005	ACALTAGE	ROD CT ASS	2 30 GN 25	AGTT HES	SERVIUSS	TOTLINES	DOCSAGES	80803865
t	4			172	19	45.345	4.573	49115	36765	11090	53.311	15 1 3	1255-
2	ь			200	21	41.755	3,506	127299	23116	11780	55.237	1104	1647
1	Œ			346	172	94.777	20.041	128522	29073	9 1 160	111.ded	2473	1228
4	ų.			17	2	2.000	0.466	3457	tet '	580	2.386	171	65 '
5	•			92	30	20.375	6.727	41706	16209	10770	75.420	1123	458
5	£			317)1	49.310	4.252	109565	35510	12310	75.193	1771	2107
7	4			418	59	76.383	5.052	110586	27117	27444	45.369	2458	27 10
. 8	h			182	7.0	45.004	9.705	14 14 76	45273	28462	67. 125	2107	2077
4	i,			216	95	44.544	8.606	134639		32569	66.266	2 160	1575
10	i			1.0	6	14.373	0.000	1 16 18	13022	11942	15:258	763	255
11	X.			45	i a	9.627	0.527	16675		5416	10.172	255	219
12	1			74	2.2	10.422	2.331	34532		6950	17.271	760	541
1.1	3			216	17	17.999	1.374	37934	1905	519	20.648	140	1274
14	q			51	;	4.959	0.110	11590		350	9.004	245	423
15	3			8.3	14	11.976	1.123	29345		5310	14.765	366	5 10
16	ų			105	16	12,227	1.571	23435		3381	14.361	527	660
17	વં			7 1	1	9.568	0.892	14023		2 14 9	14.252	511	114
18	è			7.2	29	18.580	7.338	15202		5071	12.812	973	795
19				1 1	15	2-451	1.947	7710		2010	5.497	136	10.1
20	t			18	· 5	4.150	0.000	9775		2417	9.525	169	138
21	ű			39	17	7. 150	2-049	10115		1234	2.727	284	300
22	•			23	o	2.052	0.000	5182		1118	2.052	61	135
23	ý			41	a ·	4.721	0.000	11156		1627	5.204	163	289
	•			71	•		4.556	11100	1,0,	1427	7.254	103	23.
085	2001.101	10 ACO4 85	PUZEGKS.	23K1 JVSO	PEACHES	LLLHES	EFFORT	a Lizpar	HEACTARS	20040020	SOULINSO	92440050	OTTARVSK
							24 1 4 4 1	4-4-4-		200,000	40001404	3 E 4.10 B 3 Q	
1	20 1	292	4604	46.4582	125880	136970	78.675		2056.17	361	. 21.937	29584	0.263634
2	20 1 28 J		_				78.675	85.606	2056.17		-	-	
		292	4604	46.4582	125880	136470	78.675 95.384	85.606 103.997	2056.17 1932.34	361 441	21.337	29584	0.263634
3.	28) 510 24	292 155	4604	46.4582 46.2114	125880 152615	136970	78.675 95.384	85.606 103.997	2056.17 1912.34 7179.30	361	21.037	29584 40000	0.263634
2	28) 510 24 374	292 155 587	4604 6871 11976	46.4582 46.2114 90.1568	125880 152615 157600 3440	136970 166395 200760 4420	78.675 95.184 98.500 2.400	85.606 103.997 125.475 2.762	2056.17 1932.34 7179.36 4.00	361 441 14884	21.337 12.292 401.642 0.236	29584 40000 119716	0.263634 0.219775 0.244882
3.	28) 510 24	292 155 587	4604 6871 11976 16d	46.4582 46.2114 90.1568 2.3772	125880 152615 157600 3840 57915	136970 166395 200760 4420 68705	78.675 95.184 98.500 2.400 30.197	85.606 103.997 125.475 2.762 42.741	2056.17 1912.34 7179.30 4.00 403.31	361 841 14884 4	21.337 12.292 401.642 0.236 45.253	29584 40000 119716 361 3454	0.263634 0.219775 0.244882 0.105263
2 3 8 5	28) 510 24 374	292 155 587 24 423	4604 6871 11976 164 3013	46.4582 46.2114 90.1568 2.1772 31.1440	125880 152615 157600 3440	136970 166395 200760 4420	78.675 95.184 98.500 2.400 30.197 90.672	85.606 103.997 125.475 2.762 42.741 98.166	2056.17 1912.34 7179.30 4.00 403.31 2412.07	361 %41 14884 4 700 961	21.337 12.292 401.642 0.236	29584 40000 119716 361	0.263634 0.219775 0.244862 0.105263 0.218207
2 3 9 5 6	28) 510 24 374 515	292 155 587 24 423 618	4604 6871 11976 164 3033 7500	46.4582 46.2114 90.1568 2.1772 31.1440 54.5114	125880 152615 157600 3840 57915 145075	136970 166395 200760 4420 68705 157345	78.675 95.184 98.500 2.400 30.197 90.672 89.316	85.606 103.997 125.475 2.762 42.341 98.166	2056.17 1912.34 7179.30 4.00 403.31 2412.07 5911.00	361 %41 14884 4 700 961 3481	21.837 12.292 401.642 4.236 45.253 19.380 31.945	29584 40000 119716 361 3454 113569	0.263634 0.219775 0.244882 0.105263 0.218207 0.146338
2 3 9 5 6 7	28 J 510 24 374 515 519	292 155 587 24 423 618 619	4604 6871 11976 164 3013 7500 7527	46.4582 46.2114 90.1568 2.1772 31.1440 54.5314 78.5802 49.4682	125880 152615 157600 3890 57915 145075 143705	136970 166395 200760 4420 68705 157345 171149 223211	78.675 95.384 98.500 2.400 30.197 90.672 89.316	85.696 103.997 125.475 2.762 42.341 98.166 106.968 119.507	2056.17 1932.34 7179.30 4.00 403.31 2432.07 5911.00 2025.36	361 441 14884 4 700 961 3441 4400	21.337 12.292 401.642 45.253 19.380 21.945 94.187	29584 40000 119716 361 364 113569 174724 33124	0.263634 0.219775 0.244882 0.105263 0.218207 0.146339 0.183931 0.247275
2 3 5 6 7 8	28 J 510 24 374 535 519 371	292 155 587 24 423 619 412	4604 6871 11976 164 3013 7500 7527	46.4582 46.2114 90.1588 2.1772 31.1440 54.5114 78.5802 49.4682 49.4682	125880 152615 157600 3090 57915 145075 145075 143705 194749 173967	136970 166395 200760 4420 88705 157345 171149 223211 212636	78.675 95.384 98.500 2.400 30.197 90.672 89.316 121.718	85.696 103.997 125.475 2.762 42.741 90.166 106.968 119.507	2056.17 1932.34 7179.36 4.00 403.31 2432.07 5911.00 2025.36	361 14884 4 900 961 3481 4900 4225	21.337 12.292 401.642 45.253 14.380 21.945 94.187 74.063	29584 40000 119716 361 3804 113569 174724 31124 46656	0.263634 0.219775 0.244882 0.105263 0.219207 0.146338 0.183931
2 3 5 6 7 8	281 510 24 374 535 519 371	292 355 587 24 423 618 619 412 444	4604 6871 11976 164 3013 7500 7527 15017 14561	46.4582 46.2114 90.1588 2.1772 31.1440 54.5114 78.5802 49.4682 48.9684 14.9500	125880 152515 157600 3440 57915 145075 143705 194749 173967 44660	136970 166750 260750 4420 4420 50705 157145 171149 221211 212636 56602	78.675 95.184 98.500 2.400 30.197 90.672 89.316 121.718 112.379 27.912	95.606 103.997 125.475 2.762 42.741 90.166 100.908 139.507 132.497 35.176	2056.17 1932.34 7179.30 4.00 403.31 2432.07 5911.00 2025.36 1993.39 221.21	361 941 14884 4 900 961 3491 4900 4225 0	21.937 12.292 401.642 0.236 45.253 19.380 31.945 94.187 74.183	29584 40000 119716 361 3404 113569 174724 31124 46656 4649	0.261634 0.219775 0.244682 0.105261 0.219207 0.146338 0.183931 0.247275 0.2056855 0.2056855
3 9 5 6 7 8 9	28] 510 24 374 535 519 371 391 102	292 155 587 24 423 619 412 444 113	4604 6871 11976 364 3013 7500 7527 15017 14561 1589	46.4582 46.2114 90.1568 2.1772 31.1440 54-5314 78.5802 49.4682 48.9684 14.9500 9.7360	125880 152615 157600 3840 57915 145275 143705 194749 173967 44660 21661	136970 166395 200760 4420 4420 46705 157135 171149 223211 212636 56602 27097	78.675 95.384 98.504 2.400 35.197 90.672 89.672 89.676 121.718 112.379 27.912	35.606 103.997 125.475 2.762 42.741 98.166 106.788 119.507 112.497 132.497 16.316	2056.17 1932.34 7179.30 4.00 4.00 4.33.01 2432.07 5911.00 2025.36 1993.39 221.21 92.58	361 1484 4 700 961 3441 4400 4225 0	21.337 12.292 401.542 4.236 45.253 19.380 21.945 94.187 74.063 2.300 0.279	29584 40000 119716 361 3464 111569 174724 33124 46656 2025	0.263634 0.219775 0.244882 0.105263 0.219207 0.146338 0.183931 0.247275 0.206685 0.153925 0.213933
2 3 5 6 7 8 9 10 11	28] 510 24 374 535 510 37] 391 102 55	292 155 587 24 423 618 619 412 444 113 76	4644 6871 11976 164 3013 7527 7527 15017 14561 1589 1476 2467	46.4582 46.2114 90.1568 2.1772 31.1440 54.5114 78.5802 49.4682 49.9504 14.9504 72.1118	125880 152615 157600 3080 57915 145075 143075 143705 143705 143705 143705 143705 143705 143705 143705 143705 143705 143705 14470	136970 166395 200760 4420 44705 157345 171149 223211 212636 56602 27097 53282	78.675 95.384 98.500 2.400 30.197 90.672 49.416 121.718 112.479 27.912 13.538 28.957	35.606 103.997 125.475 2.762 42.241 98.166 106.988 119.547 132.497 15.176 16.316 31.301	2056.17 1932.J4 7179.J0 4.00 403.J1 2432.07 5911.00 2025.J6 1993.J9 221.Z1 92.68 117.12	361 %41 14884 900 961 3441 4400 4225 0 64	21.337 12.292 401.642 45.253 19.380 11.945 94.187 74.063 0.279 5.334	29584 40000 119716 361 364 113569 174724 31724 46656 4649 2025 5476	0.263634 0.219775 0.244882 0.105261 0.218207 0.146338 0.148331 0.247275 0.20665 0.159925 0.213933 0.146243
2 3 5 6 7 8 9 10 11 12	28 J 51 0 24 374 535 517 371 39 I 102 55 114 26 J	292 155 587 24 423 639 432 444 113 74 161 263	4604 6871 11976 164 3013 7520 7527 7527 15017 14561 1589 1476 2467 270	46.4582 46.2114 90.1568 2.1772 31.1440 51.514 78.5802 49.7684 14.9500 9.7360 9.7360 12.1118 18.5288	125880 152615 157600 3840 57915 145075 143705 194749 173967 44660 21661 46332 39919	136370 166395 200760 4420 157365 157135 171149 223211 212636 56602 27027 53282 40178	78.675 95.184 98.500 2.400 30.197 90.672 89.416 121.718 112.379 27.912 13.538 13.538 13.538 13.538	35.606 103.997 125.475 2.762 42.741 98.166 106.968 119.507 119.507 15.176 16.36 131.301 25.236	2056.17 1912.34 7179.30 4030.01 2412.07 5911.00 2025.36 1993.39 221.21 92.58 117.12	361 941 14884 4 900 961 3441 4490 4225 0 64 484 289	21.317 12.292 401.642 40.216 45.253 11.380 31.945 94.187 74.063 3.000 6.279 5.314 1.388	29584 40000 119716 361 3904 111569 174724 33124 4656 4649 2025 5416	0.263634 0.219775 0.244882 0.105263 0.214207 0.146338 0.184931 0.247275 0.206685 0.159925 0.213933 0.146243 0.381329
2 3 8 5 6 7 8 9 10 11 12 13	20 J 51 0 24 374 515 517 191 102 55 114 26 J 73	292 155 587 24 413 619 412 444 111 161 261	4644 6871 11976 164 3013 7520 7527 75017 14581 1476 2467 270 590	46.4582 46.2114 90.1588 2.1772 31.1440 54.514 78.5842 49.4682 49.4682 47.9684 14.9500 72.1118 18.5298 5.7680	12 5880 152615 157600 3080 57915 145075 143705 173967 44661 46332 37919	136970 166395 200760 4420 44705 157145 171149 223211 212636 26602 27097 53282 40178 12660	78.625 95.184 98.500 2.400 30.197 90.672 49.416 121.718 112.479 27.912 11.518 28.957 7.594	35.606 103.997 125.475 2.762 49.166 106.908 139.507 132.497 132.497 15.176 16.936 31.301 25.236 7.712	2056.17 1912.34 7179.30 403.01 2412.07 5911.00 2025.36 1991.39 221.21 92.58 117.12 323.76 24.59	361 841 14884 900 961 3441 4900 4225 0 64 484 289	21.337 12.292 431.642 0.236 45.253 19.380 31.945 94.187 74.063 3.300 0.279 5.334 1.348 0.417	29584 40000 119716 361 3404 113549 174724 33124 46656 4649 2025 5416 46656 2601	0.263634 0.219775 0.244882 0.105263 0.219207 0.148338 0.183931 0.247275 0.206085 0.159925 0.213933 0.146243 0.381329
2 3 8 5 6 7 8 9 10 11 12 13	20 J 51 0 24 J 74 5 J 5 J 5 J 7 J 9 T 10 2 5 5 11 4 26 J 7 J 11 5	292 155 587 24 423 619 412 444 113 74 161 261 73	4604 6871 11976 164 3013 7527 7527 15017 1589 1476 2467 270 590 1741	46.4582 46.2114 90.1564 2.1772 31.1440 54.5514 78.5842 49.4682 49.7360 9.7360 9.7360 12.1118 18.5298 5.7680	125880 152615 157600 3490 57915 145075 143705 194749 173967 44660 21661 46332 39919 12310 326153	136970 166395 200760 4420 44705 157345 171149 223211 212636 56602 27097 53282 40178 12660	78.625 95.184 98.500 2.400 3.197 90.672 89.876 121.718 112.319 27.012 13.518 28.957 24.399 7.594 22.596	35.606 103.997 125.475 22.75 42.741 98.166 108.788 119.507 119.507 15.176 13.301 25.216 7.112	2056.17 1912.Ju 7179.30 4.00 403.01 5911.00 2025.J6 1993.J9 221.21 92.58 117.12 323.76 24.59	361 841 14884 4 700 961 3441 4400 4225 0 64 444 289	21.317 12.292 401.692 0.216 45.253 11.985 94.187 74.063 0.207 5.314 1.808 0.017	29584 40000 119716 361 3904 111569 174772 31122 4655 4649 2025 5416 46656 2601 6689	0.263634 0.219775 0.244882 0.105263 0.219207 0.146338 0.181931 0.247275 0.159925 0.259825 0.213933 0.146243 0.097235 0.143129
2 3 8 5 6 7 8 9 10 11 12 13 14 15 16	20 J 51 O 24 J 37 4 5 J 5 5 J 7 J 7 J 10 2 5 5 1 1 4 26 J 7 J 1 1 5 1 1 6	292 155 587 24 423 619 412 444 111 74 161 261 73 141	4604 6871 11976 164 3013 7500 7527 15017 14581 1589 1476 2467 270 590 1241 1395	46.4582 46.2114 90.1564 2.1772 31.1440 54.5314 78.5842 49.4682 49.7640 9.7360 12.1118 18.5288 5.7640 12.4554	125880 152615 157600 3480 57915 145075 143705 143705 143705 24660 21661 46312 39919 12310 36453 29427	136970 166395 200760 4420 44705 157345 171149 223211 212636 27097 53282 27097 53282 12660 41461 41461	78.675 95.184 98.500 2.400 30.197 90.672 49.416 121.718 112.379 27.912 13.538 28.957 24.389 7.594 22.596 18.192	35.606 103.997 125.475 2.762 42.741 98.166 108.988 119.547 112.497 15.176 16.936 33.301 25.236 7.712 25.236	2056.17 1912.34 7179.30 4.30 403.31 2412.07 5911.00 2025.36 1991.39 221.21 92.68 117.12 321.76 24.59 141.39	361 841 14884 700 961 3441 4400 4225 0 64 448 289 1	21.337 12.292 401.492 45.253 19.380 21.945 94.187 74.063 2.000 0.279 5.314 1.348 0.417 1.750	29584 40000 119716 361 3804 111569 1747724 33124 4656 4649 2025 5476 2601 46656 2601 11025	0.261634 0.219775 0.244882 0.105261 0.218207 0.146313 0.18391 0.247275 0.206085 0.159925 0.215933 0.146243 0.381329 0.097235 0.114448
2 3 8 5 6 7 8 9 10 11 12 13 14 15 16	20 J 51 O 24 J 74 SJS 51 O J 7 J J 9 T 10 Q 55 T T 4 26 J 7 J 1 T 5 1 L 6 1 O Q	292 155 587 24 423 619 412 444 113 74 161 261 73 143 140 113	4604 6871 11976 164 1013 7500 75017 14561 1589 1476 2467 270 590 1241 1195	46.4582 46.2114 90.1564 2.1772 31.1440 78.5942 49.9684 14.9504 79.7360 9.7360 9.7360 12.4554 12.7542	125880 152615 157600 3440 57915 145075 194749 173967 44660 21661 46312 39919 12310 36153 29427	136970 166395 200760 4420 64705 157145 221211 212634 56602 27027 53282 40178 12660 91461 32808 17908	78.625 95.184 98.500 2.400 30.197 90.672 49.416 112.379 27.912 11.518 28.957 24.349 7.594 22.596 18.192	85.606 101.997 125.475 2.762 42.741 98.166 108.788 1132.497 35.176 16.36 113.301 25.216 7.712 25.219 20.505	2056-17 1912-Ju 7179-30 4-30 4-31-37 2412-07 5911-00 2025-36 1991-39 221-21 92-58 117-12 321-26 24-59 141-39 149-50 91-55	361 \$41 14884 900 961 3441 4900 4225 0 64 444 289 1 1966 49	21.337 12.292 401.692 0.236 45.253 19.380 31.995 94.187 74.063 3.000 0.279 5.314 1.398 0.417 1.750 2.366	29584 40000 119716 361 3904 111569 174724 31122 46656 4649 2025 5476 46656 2601 6889 11025 5001	0.261614 0.219775 0.244882 0.105261 0.219207 0.146318 0.181931 0.241275 0.206085 0.206085 0.16243 0.184243 0.184243 0.116449 0.116449
2 3 9 56 7 8 9 10 11 12 13 14 15 16 17 18	28 J 510 24 374 515 517 171 102 55 114 26 J 115 116 100 148	29 2 25 5 24 42 3 61 9 41 2 44 4 11 1 16 1 26 1 7 3 14 1 16 1 26 1 17 3 14 1 18 0 18 1 18 0 18 1 18 0 18 1 18 0 18 1 18 0 18 1 18 0 18 0	4604 6871 11976 364 3013 7500 7527 15017 14561 1589 1476 2467 270 590 1241 1395 1151 2354	46.4582 46.2114 90.1584 2.1772 31.1440 54.5314 78.5802 49.4682 49.4682 49.7660 12.1118 5.7660 12.4554 12.4554 12.4554 12.7542 10.51084	125880 152515 15760 3440 57915 145075 194749 173967 44660 21661 46332 39419 12310 36453 29427 1552	136970 166395 200764 4420 4420 44705 157145 171149 223211 212636 27097 53282 40178 12660 179072	78.675 95.184 98.500 2.400 30.197 90.672 89.416 121.718 112.319 27.912 11.518 28.957 7.594 18.192 9.724	35.606 101.997 125.475 2.762 42.241 98.166 108.988 119.507 112.497 135.176 16.36 131.301 25.216 7.712 25.219 20.505 11.192 30.457	2056.17 1912.34 7179.30 4.30 4.30 2412.07 5911.00 2025.36 1993.39 221.21 92.58 117.12 321.76 24.59 141.39 141.39 149.50 91.55	361 841 1484 900 961 3441 4900 4225 0 64 484 289 1 196 256 49	21.337 12.292 431.642 4.216 45.253 19.380 31.945 94.187 73.063 3.300 0.277 5.334 1.398 0.417 1.750 0.417 1.750 0.417	29584 40000 119716 361 3404 111569 174724 31124 46656 46646 2025 5476 2601 1025 50416 11025 50416	0.261634 0.219775 0.244882 0.105261 0.218207 0.146338 0.18338 0.18338 0.247275 0.25085 0.159925 0.159925 0.159925 0.159925 0.1831329 0.097735 0.143108 0.118448 0.118448 0.134761 0.2574444
2 3 8 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19	28 J 510 24 374 515 517 J71 J91 102 55 114 26 J 73 115 116 100 148 38	292 155 587 24 423 619 412 444 111 74 161 261 73 141 180 113 245 29	4604 6871 11976 1664 3013 7520 7527 75217 18361 1589 1476 2467 270 721 1395 1151 2354	46.4582 46.2114 90.1564 2.1772 31.1440 78.5802 49.4684 74.9504 74.9504 77.160 12.1118 18.5298 5.7680 12.4554 12.7542 10.5108 21.5084 31.5084	125880 152615 157600 3440 57915 145075 145075 143705 173767 44660 21661 46332 39919 12310 36453 29427 15559 44293	136970 166395 20076 4420 44705 157145 171149 223211 212636 56602 27097 53282 40178 12660 41461 32808 17908 49372	78.625 95.184 98.500 2.400 30-197 90.672 89.316 121.718 112.379 27.312 112.379 27.312 12.357 24.369 7.594 22.596 18.157 24.768 19.724 27.08 10.281	35.606 101.997 125.475 2.762 42.741 98.166 106.908 119.307 112.497 15.176 16.936 11.201 25.216 26.505 11.192 30.357 7.517	2056-17 1912-Ju 7179-JO 4-JO 4-JO 4-JO 4-JO 75911-GO 2025-J6 1991-J9 221-21 92-58 117-12 J2J-76 24-59 141-J9 149-50 91-55 J48-J4	361 \$41 14884 4 900 961 3441 4400 4225 0 64 444 289 1 1 196 256 49 341 225	21.337 12.292 401.642 0.236 45.253 19.380 31.945 94.187 74.063 3.000 0.279 5.314 1.358 0.417 1.750 2.464 0.776 61.414	29584 40000 119716 361 360 113569 174724 31124 46656 4649 2025 5476 46656 2601 6489 11025 5041 5184	0.261634 0.219775 0.244882 0.105261 0.218207 0.146338 0.181931 0.241275 0.159925 0.159925 0.159925 0.146243 0.381329 0.097235 0.14448 0.134761 0.2514448
2 3 5 5 7 8 9 10 11 12 13 14 15 16 17 18 19 20	28 J 510 24 574 515 517 102 55 114 26 J 115 116 148 38 41	29 2 25 5 24 4 23 6 36 19 4 44 4 11 3 16 1 26 1 18 0 11 1 18 0 11 1 24 5 24 9 44 4	4604 6871 1976 1664 3033 7527 15017 1509 1476 2467 2467 570 570 1243 1354 1354 1354	46.4582 46.2114 90.1564 2.1772 31.1440 54.5314 78.5802 49.4682 49.4682 49.4682 47.360 12.118 5.7640 12.2118 5.7640 12.4554 12.7542 10.51084 1.0602 4.2330	125880 152615 157600 3480 57915 145075 143705 143709 173967 446312 46312 39919 12310 29427 15553 160049 13221	136970 166195 20074420 44420 44705 1571345 171149 221211 212636 27097 53282 27097 12660 1790178 12660 1790178 12660 1790178 12660 1790178 12660 1790178 12660 1790178 12660 1790178 12660 1790178 12660 1790178 12660 1790178 12660 1790178 17660 1766	78.675 95. J84 98. S00 2. 400 J0. 197 90.672 89. 416 121.718 112. 379 27. 912 113.518 28.957 24.389 7.594 18.192 9.724 27.681 5.281 5.281	35.606 103.997 125.475 2.762 42.741 98.166 104.798 119.507 119.507 119.307 25.216 7.712 25.218 20.505 11.192 30.457 7.517 9.774	2056.17 1912.Ju 7179.JO 403.J1 2412.07 5911.d0 2025.J6 1991.J9 221.21 92.58 117.12 327.76 24.59 141.J9 141.J9 149.50 91.55 348.Ju 6.01 17.J1	361 941 14884 961 3491 4490 4225 0 644 484 289 1 196 256 49	21.337 12.292 441.642 4.253 19.380 11.945 94.187 74.063 3.000 0.277 5.314 1.550 2.463 0.017 1.1550 2.463 0.796 61.414 3.791	29584 40000 119716 361 3404 111569 174724 31124 46656 4669 2025 5476 2601 6689 11025 5041 16125 1614 1614	0.261634 0.219775 0.244882 0.105261 0.218207 0.146338 0.18338 0.247275 0.20685 0.159925 0.159925 0.1931329 0.097735 0.14448 0.134761 0.158538 0.143108
23 m 56 7 a 9 0 11 12 13 15 6 17 8 9 0 2 1	28 J 510 24 J 74 J 515 517 J 10 J 10 J 11 J 10 J 11 J 10 G 14 B 18 B 18 B 18 B 18 B 18 B 18 B 18 B 18	292 1557 24 423 619 412 444 111 74 161 261 73 141 113 245 144 115	4604 6871 11976 166 3013 7500 7527 7501 1589 1476 2467 270 7590 1241 1195 1151 2154 465 856	46.4582 46.2114 90.1564 2.1772 31.1440 54.5914 78.5802 49.7864 14.9504 9.7360 12.1118 18.5298 5.7680 12.4554 10.5188 21.5084 1.5	125880 152615 157600 3440 57915 145075 143705 173767 44660 21661 46332 39913 36153 29427 15253 44529 44529 11221	136970 166395 20076 4420 44705 157145 171149 223171 212636 56602 27097 53282 40177 12660 41463 17908 17908 19372 12639 17908	78.675 95.184 98.500 2.400 30.197 90.672 89.316 112.379 27.312 112.379 27.312 28.357 24.399 22.596 22.596 22.596 27.681 8.261 9.1281	85.606 101.997 125.476 2.762 42.741 98.109 106.988 119.507 15.176 16.936 31.301 25.210 25.210 20.305 11.192 30.357 7.517 9.773	2056.17 1912.Ju 7179.JO 403.J1 2412.07 5911.00 224.5.J6 1993.J9 221.Z1 92.58 117.72 J21.59 141.J9 149.50 91.55 J48.54 6.01 17.J1 54.J2	361 841 1484 4 900 961 3441 4400 4225 0 64 444 289 1 1 196 256 49 841 225	21.337 12.292 431.642 0.236 45.253 19.380 31.945 94.187 74.063 3.300 0.279 5.334 1.346 0.417 1.750 0.796 61.414 3.791 0.000	29584 40000 119716 361 3901 113549 174724 31124 46656 4649 2025 5476 46656 4666 2601 6489 11025 5041 5184 1694 1694	0.261634 0.219775 0.244882 0.105261 0.218207 0.146313 0.183931 0.201275 0.206085 0.159925 0.159925 0.1164283 0.181329 0.97235 0.1413108 0.116448 0.138761 0.138761 0.138761 0.138761 0.138761
2 3 5 5 7 8 9 10 11 12 13 14 15 16 17 18 19 20	28 J 510 24 574 515 517 102 55 114 26 J 115 116 148 38 41	29 2 25 5 24 4 23 6 36 19 4 44 4 11 3 16 1 26 1 18 0 11 1 18 0 11 1 24 5 24 9 44 4	4604 6871 1976 1664 3033 7527 15017 1509 1476 2467 2467 570 570 1243 1354 1354 1354	46.4582 46.2114 90.1564 2.1772 31.1440 54.5314 78.5802 49.4682 49.4682 49.4682 47.360 12.118 5.7640 12.2118 5.7640 12.4554 12.7542 10.51084 1.0602 4.2330	125880 152615 157600 3480 57915 145075 143705 143709 173967 446312 46312 39919 12310 29427 15553 160049 13221	136970 166195 20074420 44420 44705 1571345 171149 221211 212636 27097 53282 27097 12660 1790178 12660 1790178 12660 1790178 12660 1790178 12660 1790178 12660 1790178 12660 1790178 12660 1790178 12660 1790178 12660 1790178 12660 1790178 17660 1766	78.675 95. J84 98. S00 2. 400 J0. 197 90.672 89. 416 121.718 112. 379 27. 912 113.518 28.957 24.389 7.594 18.192 9.724 27.681 5.281 5.281	35.606 103.997 125.475 2.762 42.741 98.166 104.798 119.507 119.507 119.307 25.216 7.712 25.218 20.505 11.192 30.457 7.517 9.774	2056.17 1912.Ju 7179.JO 403.J1 2412.07 5911.d0 2025.J6 1991.J9 221.21 92.58 117.12 327.76 24.59 141.J9 141.J9 149.50 91.55 348.Ju 6.01 17.J1	361 941 14884 961 3491 4490 4225 0 644 484 289 1 196 256 49	21.337 12.292 441.642 4.253 19.380 11.945 94.187 74.063 3.000 0.277 5.314 1.550 2.463 0.017 1.1550 2.463 0.796 61.414 3.791	29584 40000 119716 361 3404 111569 174724 31124 46656 4669 2025 5476 2601 6689 11025 5041 16125 1614 1614	0.261634 0.219775 0.244882 0.105261 0.218207 0.146338 0.18338 0.247275 0.20685 0.159925 0.159925 0.1931329 0.097735 0.14448 0.134761 0.158538 0.143108

Data Set 1

					9012075	Atsii naos	1 50K TOT	fr rigss (F C002				
085	721159			2005 K3 K	ZOOLODE	*EAFTRE2	ROOLINE	S 6807488	S AGITHRS	SERTRES	ZSKLITOT	00061655	ROBCHEGS
1	d.			12	2	2.000	0.486	3457	383	580	2.886	17.1	45
2	1			93	q	14.873	0.000	31638	1 10 22	11942	15.258	763	255
1	k.			45	a	9.627	0.527	16675	òB₽U	5436	10.172	255	219
.4	L			. 74	2.2	10.322	2.331	34532	11800	6950	17.271	760	541
5	4			216	17	17.999	1.174	379)4	1905	\$39	20.548	140	1274
6	2			5.1	1	4,959	0.134	11590	720	350	9.004	245	423
7	ف			83	Į ų	11.876	1. 12 1	29 38 5	6748	5310	14.765	Ján	5 10
8	o			105	16	12-227	1.571	21015	6392	3 3 8 1	14.861	527	680
9	à			7.1	7	9.568	0.392	14023	1535	2349	14.262	511	311
10	ė			72	2 7	18.680	7.83A	35202	9091	5079	12.922	973	795
1.1	4			13	15	2.451	1,947	7780	2269	2010	5.497	130	103
1.2	ŧ			38	٥	4.160	0.000	9775	3446	2417	4.525	169	158
13	ū			39	17	7.150	2.049	10115	1290	1284	9.727	244	100
14	*			23	. 0	2.052	0.004	5182	2290	1118	2.052	61	115
15	4			4.1	á	4.921	0.000	11166	1765	1627	5.204	16 3	294
085	20017005	X63C0462	MUARUES	OSTLINES	29.5GU8S	ALLHRS	EFFORT	ALLETST	PERTJUSE	aconacsq	RODLINSQ	#84300SQ	COTABER
1	24	24	168	2.1772	1840	4420	2.4000	2, 7625	9.000	4	0.2362	361	0.105263
2	10.2	111	1589	14.9500	44660	56602	27.9125	15. 3762	221.206	0	0.0000	8649	0.159925
3	35	74	1476	9.7360	21661	27097	13.5381	16.7156	92.679	64	0.2777	2025	0.213913
4	114	161	2467	12.1118	44332	5 3 2 8 2	28.9575	33. 10 12	117.116	4 4 4	5.4316	5476	0.146243
5	26.1	26 J	270	18.5288	33839	40)78	24.8994	25.2362	121.964	299	1.8479	46656	0.091129
6	73	73	590	5.7640	12310	12660	7.6937	7,9125	24.592		0.0169	2541	0.397235
7	115	143	1283	12.4554	36153	41461	22.5956	25.9144	141.007	196	1.7543	6499	g. tulige
Ħ	110	180	1195	12.7542	29427	32808	18.3919	20.5050	149.500	256	2.9680	11025	0.115448
9	100	113	1151	10.5108	15559	17908	9.7294	11. 1925	91.547	49	0.7957	5041	0.134761
10	144	245	2354	21.5084	44293	49372	27.6611	30.3575	340.742	841	61,4342	5184	0.259444
11	3 4	19	312	3.0602	10049	12059	6.2806	7.5169	6.007	225	3.7909	169	0.198534
12	41	44	465	4.2330	1 122 1	15438	8.2631	9.7737	17.306	0	0.0000	1464	0.109474
13	43	114	856	7.3254	11405	12689	7.1291	7.9306	54.022	289	4. 1984	1521	0.188462
14	23	35	221	2,0520	7472	8590	4.6700	5. 16 87	4.211	0	0,000	529	0.089217
15	4.8	7 %	50.6	4.9776	12911	14558	8-0519	9. 3987	7a. 21A	4	0.0000	1681	D. 12602 N

PROJECTS WITH SOR OR HORE TOTAL LIBES OF CODE

085	127789			# EA4002	2001001	MERFERE	40061465	2890845	AGATHES	25 B A B 52	tortines	00CP &G 85	MU 4CHMCZ
1	1			172	19	45.145	4.673	89115	36765	11090	50.911	16.13	1255
Z	b			204	21	43,955	3,506	129299	23316	11780	55.237	1104	1049
3	~			346	122	84.729	20.041	128522	29078	4 3 16 0	111,468	2471	1229
4	ŧ			337	11	49.316	4.252	109565	35510	12310	75.193	179]	2107
5	3			418	59	76.341	5.652	116586	27119	27944	85.169	2458	2710
6	b.			182	70	45.304	9.705	149475	45273	28462	67.125	2107	2017
7	Ĺ			216	65	46, 545	3.606	134639	45328	12669	6a. 266	2360	1575
a	ı			*09	84	61.950	14.297	123143	26078	19265	99.513	2695	2761
085	1073005	10400475	aga bon 2	DEVLIMES	28 52 58 9	LLLSSS	EFFORT	ALLEPBT	aga CI a S O	ÇZQOBGOS	MG0L1MSQ	*EWAGGSQ	STRARTE
Ţ	20 t	292	4634	46.4582	125880	136970	78.675	8505	2056.17	361	21-837	29584	0.263634
2	24.3	155	6671	46.2114	152615	166 195	95. 184		1932.34	941	12.292	40000	0.219775
3	510	587	11976	90.1568	157600	200160	98.500		7179.30	14894	101.642	119716	0.244882
9	535	618	7500	54.5314	145075	157385	90.672		29 12 . 07	961	19.380	111569	0.146138
5	519	639	7527	78.5802	141705	171149	69.916	106.768	5711.30	3481	31.945	174729	0.191911
6	373	432	15317	49.4682	194749	223211	121.718	139.507	2025.16	4900	94.187	33124	0.247275
7	391	444	19561	48.9664	179967	212636	112.479		1993-29	+225	74.063	46656	0.206985
ð	604	451	7379	57,4626	151221	179486	94.513		3837.30	7054	204.404	167281	0. 151467



SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION	PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
TR-1520		
I. TITLE (and Subtitle)		5. TYPE OF REPORT & PERIOD COVERED
Finding Relationshi		
Effort and Other Variables in the	SEL	
		6. PERFORMING ORG. REPORT NUMBER
		8. CONTRACT OR GRANT NUMBER(s)
V AUTHOR(s) Victor R. Basili		
N. Monina Panlilio-Yap		NASA NSG-5123
PERFORMING ORGANIZATION NAME AND ADDRESS		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Department of Computer Science		AREA & WORK UNIT NOMBERS
University of Maryland		
College Park, MD 20742		
1. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE
NASA/Goddard Space Flight Center		July 1985
Greenbelt, Maryland 20771		13. NUMBER OF PAGES 31
4. MONITORING AGENCY NAME & ADDRESS(If different	from Controlling Office)	15. SECURITY CLASS. (of this report)
		IDIOI AGGINIDO
		UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
7. DISTRIBUTION STATEMENT (of the abstract entered i	n Block 20, if different from	m Report)
,, olo () (o) () (o) () (o) () (o) () (o) () (o) () (o) () (o) () (o) () (o) (o) () (o) (
		<u> </u>
3. SUPPLEMENTARY NOTES		
	•	•
. KEY WORDS (Continue on reverse side if necessary and	identify by block number)	
software development, resource est	imation	
effort models, stepwise regression		
errore moders, stepwise regression	•	
	•	
. ABSTRACT (Continue on reverse side if necessary and	Identify by block number	Estimating the amount of
effort required for a software deve	lopment project	is one of the major aspects
of resource estimation for that pro	ject. In this s	tudy, we examined the rela-
tionship between effort and other v	ariables for 23	Software Engineering Laborato
projects that were developed for NA	SA/Goddard. The	se variables fell into two
ategories: those which can be det	ermined in the e	arly stages of project de-
velopment and may therefore be usef	ul in a baseline	equation for predicting effo
n future projects, and those which	can be used mai	nly to characterize or evalu
		ing of the software developme

READ INSTRUCTIONS

