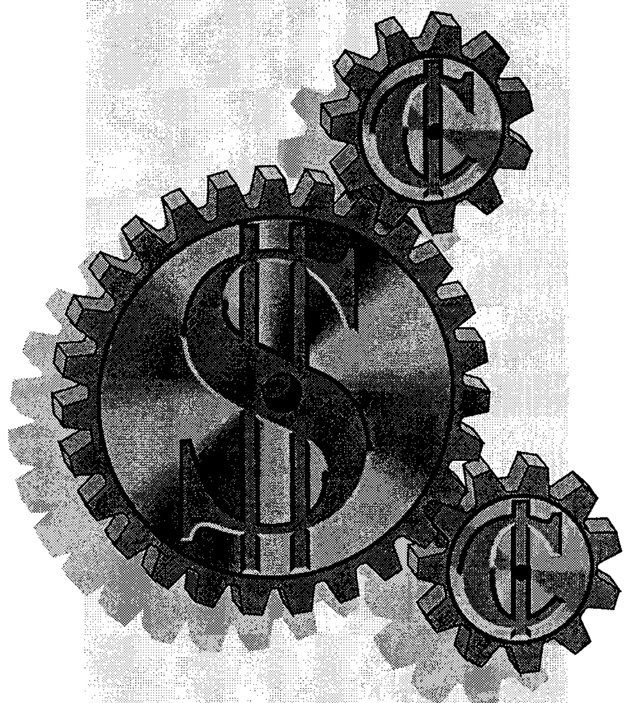
FINAL REPORT

A REVIEW AND UPDATE OF THE VIRGINIA DEPARTMENT OF TRANSPORTATION CASH FLOW FORECASTING MODEL



JAMES S. GILLESPIE
Research Scientist

CHERIE A. KYTE
Research Assistant

V·I·R·G·I·N·I·A
VTRC
TRANSPORTATION RESEARCH COUNCIL

VIRGINIA TRANSPORTATION RESEARCH COUNCIL

Standard Title Page - Report on State Project

Report No.	Report Date	No. Pages	Type Report:	Project No.:
VTRC97-R2	August 1996	49	Period Covered:	9124-030-940
				Contract No.:

Title and Subtitle

A Review and Update of the Virginia Department of Transportation's Cash Flow Forecasting Model

Key Words

cash flow
Z-vectors
regression equation
forecasting

Authors James S. Gillespie and Cherie A. Kyte

Performing Organization Name and Address:

Virginia Transportation Research Council
530 Edgemont Road
Charlottesville, VA 22903

Sponsoring Agencies' Name and Addresses

Virginia Department of Transportation
1401 E. Broad Street
Richmond, VA 23219

Supplementary Notes

Abstract

This report details the research done to review and update components of the VDOT cash flow forecasting model. Specifically, the study updated the monthly factors submodel used to predict payments on construction contracts. For the other submodel reviewed in this study, the research culminated in the formulation of a new forecasting technique for maintenance expenditures.

The report is divided into three sections. The research undertaken to update the "Z-Vectors" of the monthly factors submodel and how these new vectors may be used for forecasting purposes is presented in the first section. The development of a new model for predicting maintenance expenditures is documented in the second section of the report. How this new forecasting procedure is used is also explained in this section. Finally, recommendations are suggested to improve the Virginia Department of Transportation's cash flow forecasting techniques.

FINAL REPORT DRAFT

A REVIEW AND UPDATE OF THE VIRGINIA DEPARTMENT OF TRANSPORTATION CASH FLOW FORECASTING MODEL

James S. Gillespie
Research Scientist

Cherie A. Kyte
Research Assistant

Virginia Transportation Research Council
(A Cooperative Organization Sponsored Jointly by the Virginia Department of Transportation and the University of Virginia)

Charlottesville, Virginia

August 1996
VTRC 97-R2

Executive Summary

The Virginia Department of Transportation's Financial Planning and Debt Management Office (FPDM) uses a cash flow forecasting model to predict operations expenditures by month. Components of this general forecasting model estimate line items in the Virginia Department of Transportation (VDOT) budget. The cash flow model was developed in the early 1980's and has not been updated since 1988. This research reviewed and updated several components of the cash flow model, namely, the monthly factors submodel used to predict monthly payments to construction contracts, and the maintenance expenditures submodel used to predict maintenance payments.

The research produced an update of the monthly factors submodel, mainly by the same basic methodology used to develop the general cash flow model. However, the update used data from a seven year period, rather than from a three year period as the 1988 update did. Recalibration of the monthly factors submodel meant the re-estimation of seven sets of parameters, labeled Z-Vectors in the submodel's BASIC program.

The limited historical data make it virtually impossible to use past figures -- either the FY 86-88 data or the FY 1988-95 data -- to compare the performance of the current and the updated Z-Vectors. The FPDM should continue to generate forecasts of construction contract payouts and the "Highway System Acquisition and Construction" line item with the current monthly factors submodel, and begin to generate simultaneous forecasts by a submodel that uses a 50:50 weighted average of the FY 86-88 Z-Vectors and the FY 89-95 vectors. This will produce two parallel series of forecasts, which can then be directly compared.

This study produced a forecasting submodel for estimating maintenance expenditures using a new approach that differs in some particulars from the current FPDM model. The current submodel relies on historical averages to predict future maintenance expenditures. The new submodel relies on historical patterns plus a group of leading indicators that prefigure deviations from the historical averages. This conditional forecasting model suggests that maintenance expenditures may be influenced by seasonal patterns, nonagricultural wages and salary employment, the prime rate, and the VDOT biennial budget cycle. During the construction of the new model, the authors examined over 60 different variables in hundreds of formulations to find the forecasting equation that best fit the FY 89 through FY 95 data.

The current maintenance disbursements submodel produced an average absolute forecast error of \$3.774 million over the period FY 90 through FY 95. The new submodel produced an average absolute forecast error of \$1.392 million over the same period. It is recommended that FPDM begin to use the revised maintenance expenditures submodel to forecast maintenance disbursements.

It is also recommended that VTRC, in cooperation with FPDM, continue to monitor the performance of the updated submodels, and review the components of the cash flow forecasting model again in three years.

FINAL REPORT

A REVIEW AND UPDATE OF THE VIRGINIA DEPARTMENT OF TRANSPORTATION'S CASH FLOW FORECASTING MODEL

James S. Gillespie and Cherie A. Kyte

INTRODUCTION

The ability to forecast cash receipts and disbursements is important for efficient cash management. A 1983 study by the Virginia Highway and Transportation Research Council (now the Virginia Transportation Research Council, VTRC) found that before the 1980's the Virginia Department of Highways and Transportation (now the Virginia Department of Transportation, VDOT) experienced relatively stable cash flows and was therefore able to avoid inconvenient shortages or surpluses of cash without using sophisticated forecasting techniques. In the late 70s and early 80s, however, the Virginia Department of Highways and Transportation found that revenues and outlays, especially construction payouts, became increasingly volatile.[1] The study noted that "frustrating delays in the work program and last minute changes in ad dates can occur when cash inflows are inadequate to pay for ongoing as well as scheduled projects, as happened during the latter half of 1980. Alternatively, unnecessary delays and missed opportunities can occur when cash balances turn out to be larger than needed, as was the case during much of 1982."[2]

The Allen and Shapanka studies of 1983 and 1985 recommended new methods, termed "submodels," for predicting VDOT expenditures and revenues. For example, the authors developed a monthly factors submodel to forecast the amount of money VDOT pays out each month for construction contracts. Shortly thereafter, a consultant amalgamated the output of all the submodels into a general cash flow model. The cash flow model provides forecasts of all VDOT funds by activity. The model can provide forecasts for several years into the future, employing data on state and federal revenues to predict inflows of cash and employing data from other sources to predict the outflows. The Commonwealth Accounting and Reporting System (CARS), Department of Motor Vehicles (DMV), and various taxation figures provide additional data for the cash flow model on such items as motor vehicles sales and use tax and fuel tax revenue.

VDHT implemented the recommendations of the 1983 VTRC study in 1984 and those of the private consultant in 1986. VTRC revised the cash flow model in 1988 using 1986-1988 data, and VDOT has employed the model since that time.

---

[1] *A System for Forecasting and Monitoring Cash Flow: Phase I: Forecasting Payments on Construction Contracts*, Gary Allen and Adele Shapanka, 1983, p.2.

[2] Ibid, p.3.

PURPOSE AND SCOPE

Except for a small revision in 1988, VDOT has used the same cash flow model to forecast revenues and expenditures since 1986. FPDM requested an update of the methods currently used to predict VDOT's monthly payouts to construction and maintenance activities. The purpose of this project was to recommend revisions to the Monthly Factors Submodel which forecasts part of the "Highway System Acquisition and Construction" line item in the VDOT budget, and the Maintenance Expenditures Submodel which forecasts the "Highway System Maintenance" line item in the budget, to reflect the structural changes in VDOT operations since 1988 and produce superior forecasts.

METHOD

VTRC reviewed and updated the Monthly Factors Submodel used to forecast monthly construction contract payouts, and the Maintenance Expenditures Submodel used to forecast monthly maintenance disbursements. The two payout streams are modeled separately. The review and update proceeded in several steps.

First, available background information concerning the general cash flow model was collected and reviewed by studying the 1983 and 1985 reports that launched the Cash Flow Model, and holding conversations with VDOT personnel familiar with the models, principally Dr. Gary Allen of VTRC and Mr. Ernest Miller and Ms. Diane Mosby of FPDM.

Second, the structures of both the Monthly Factors Submodel and the Maintenance Expenditures Submodel, and the methods previously used to calibrate them, were examined in light of the data currently available. As the forecasting principles on which the maintenance forecast is based are rather simple, and as no exact mathematical form for the Maintenance Expenditures Submodel is on record, it was decided to experiment with several alternative forms for this submodel to recommend a formula that best "fits" the expenditure pattern observed in FY 1989-95.

Third, the availability and quality of the data that would be needed to revise the forecasting submodels were assessed. This assessment entailed interviewing personnel in VDOT's Construction (CD), Fiscal (FD), and Information Systems (ISD) Divisions who have access to receipt and disbursement records, and obtaining their assistance in retrieving the necessary data.

Fourth, new FY 89 through FY 95 contract data gathered with the cooperation of FPDM, CD, FD, and ISD, and the statistical analysis tools available in the Microsoft Excel spreadsheet made possible a re-estimation of the parameters of the Monthly Factors Model. For the Maintenance Expenditures Submodel, new FY 1989-95 maintenance data plus the same spreadsheet software permitted the testing of a host of explanatory factors in a variety of forecasting equations. The FPDM staff responsible for the forecasts maintained a dialogue with the VTRC research team throughout the project. FPDM assisted in identifying key members of the CD, FD, and ISD during the search for data.

Fifth, FPDM received a preliminary report document, on which they provided comment, and provisional statistical results, which they applied to their forecasts on a trial basis.

Sixth, at the conclusion of the project, the new parameter estimates for the monthly factors submodel and the most successful of the equations tested for the maintenance disbursements submodel were recommended to FPDM.

RESULTS AND DISCUSSION

Monthly Factors Submodel

The first task involved a review of the current model's components or submodels. Interviews with VDOT staff in the Financial Planning and Debt Management Division (FPDM), Fiscal Division (FD), and the Construction Division (CD) who are familiar with the cash flow model and an examination of previous documentation produced background information.[3] The monthly factors submodel produces estimates of total monthly payout to construction contracts.

Data Requirements

The monthly factors submodel required data on monthly construction contract payouts by contract number from 1985 to 1995 and data on large samples of individual contracts. The data was provided by the FD, FPDM, CD, and ISD.[4] The ISD provided five self-extracting files from the Bid Analysis Monitoring System (BAMS) and the Current Project Data from the FD. The files contained information on construction contract payouts, specifically:

1) Data on estimated contract amount, number of projects in the contract, number of bids, vendor, estimated completion date, let date, award date, work begin date, actual completion date, federal award amount, federal final amount, award amount, final contract

---

[3]*A System for Forecasting and Monitoring Cash Flow: Phase I: Forecasting Payments on Construction Contracts*, Adele Shapanka and Gary R. Allen, Virginia Transportation Research Council, 1983, and *A System for Forecasting and Monitoring Cash Flow: Phase II: Forecasting Federal and State Revenues, Contracts, Other Expenditures, and Cash Balances*, Adele Shapanka and Gary R. Allen, Virginia Transportation Research Council, 1985.

[4]Jim Hewitt, Miriam Daughtry, Carolyn Payne, David Reid, and Greg Whirley of the Finance Division, Ernie Miller, and Diane Mosby of the Financial Planning and Debt Management Division, Buddy Edwards, David Nestor, and William McDowell of Construction Division, and Tom Hutton of Information Services Division assisted in the retrieval of this data.

amount, contract payment amount from BAMS, estimated days to completion, final payment, liquidated damages, and workorders.

2) Project specifics such as type (urban, rural, primary, secondary, interstate), length, etc.
3) Data by contract number on project amounts and dates of payments.
4) Vendor names and amounts paid to them.
5) Payouts to projects by month and project number.

The FD provided a copy of its Contractor Payment Processing System (CPPS), which is a computer program that contains actual payment streams to projects and contractors and can be used to generate yearly, monthly, daily, supervisor daily, active projects, project summary, final payment, contractor, and escrow retainages reports.

Modeling and Estimation Methods

The monthly factors submodel depends on seven "Z-Vectors," which are columns of numbers used in conjunction with detailed contract data to predict payouts to construction contracts.[5] VTRC re-examined and updated these vectors.

The appearance, function, original calculation method used by Allen and Shapanka, and the updated calculations of each of the vectors are discussed below.[6] For the most part, the update of the vectors followed the methodology used by Gary Allen and Adele Shapanka in the development of the monthly factors submodel.[7] However, the update used more data and encompassed a longer period of time than the original study.[8]

ZA Vector (Final Project Cost)

- This vector is used in conjunction with initial contract estimates to predict the final cost of a project. The vector has nine entries that correspond to the nine contract size groups defined in the BASIC program.[9]
- The update of this vector was based on the original method used by Allen and Shapanka in 1983 and a later update by Allen in 1988.

---

[5] This lengthy procedure is automated in a BASIC program. Explanatory notes were added to the program in the fall of 1994.

[6] The calculation of the Z-Vectors and how they are used in the submodel are presented schematically in Appendix B.

[7] Ibid.

[8] The original study did not adjust for inflation and neither did the update since an inflation adjustment would not affect the results.

[9] The BASIC program defines nine contract size groups: (1)≤$300,000, (2)≤$500,000, (3)≤$750,000, (4)≤$1,000,000, (5)≤$2,500,000, (6)≤$5,000,000, (7)≤$8,000,000, (8)≤$10,000,000, and (9)≥$10,000,000. The program categorizes individual contracts into these groups for further analysis, e.g., to determine the amount paid out to a contract by its completion date.

- The calculation of the new ZA vector required data on construction contracts, such as contract number and estimated completion date, all sorted by contract size.[10] This data was available on disk for the calendar years 1985 through 1994.[11]
- The above databases were adjusted to remove projects that were incomplete. This left a list of contracts in each contract size group. The fraction of cost overrun for each contract was calculated by dividing the actual final contract amount by the estimated final amount.[12]
- Blank cells, zeros, and unfinished projects were removed from the sample. There were some instances where the overruns were unduly large or small and these "outliers" appeared to skew the results. To alleviate the effect of the outliers, contracts whose overrun ratios were greater than 2.0 or less than 0.5 were removed from the database. The removal of these outliers does lead to an underestimation of the overrun ratio variance in each contract size group, but since the mean is of paramount interest, this is relatively inconsequential. After this procedure, the mean may be said to be insulated from the influence of anomalous events. The number of sample contracts within each of the nine groups varied from a low of 34 to a high of 1,580 contracts.
- The column of individual percent cost overruns was averaged to determine the mean percent cost overrun for each contract size group.

ZM Vector (Monthly Payout Percent)

- The ZM Vector is used to predict the percentage of annual payout that (typically) will occur each month. The twelve vector entries correspond to the twelve months of the year. The entries are decimals that sum to one.
- Allen and Shapanka estimated the ZM Vector by using unweighted average annual shares.
- The initial approach in the update used data on monthly payouts to projects, which was supplied by ISD. The data series spanned FY 86 through FY 94.
- Each month's share of annual payout was determined, and then averaged across months. A weighted average was obtained by applying the total expenditure for each fiscal year to the monthly amount in the corresponding year. However, this method of updating the ZM Vector was rejected after consultations with FPDM.
- As an alternate method, data from CPPS were used in the re-calculation of an updated vector. This data set encompassed the fiscal years 1986 through 1994. The monthly fractions were determined for each month and then averaged over the eight fiscal years. A weighted average based on the number of payments occurring

---

[10] Other data included estimated final amount (a Fiscal Division forecast), date completed, estimated days until completion, and the actual final contract amount.

[11] Dbase 4 was employed to sort the data file by contract size and display the estimated contract size and actual payment for each contract.

[12] Sample sizes for the ZA Vector, ZD Vector, and ZC Vector are available in Appendix B.

each month was applied to the monthly fractions.[13] FY 86 was dropped from the sample payments because there were very few payments made in this fiscal year. The researchers compared the ratios of samples to populations in each contract size group to determine if the samples used were representative of the population. The sample sizes were judged to be representative.[14]

- The next step was to normalize these results. Normalization was accomplished by summing the column of fractions and then dividing each entry by the total, producing weights that sum to one.

ZF, ZS, ZL Vectors (First, Second, and Last Payouts)

- These three vectors represent the percentages of the total payments to contractors that comprise the first, second, and final payments. The vectors are used to predict the portion of contracts paid out in the first, second, and last month of a project. Each of the vectors has nine entries to correspond to the nine contract size groups.
- The update of these vectors was based on the original method used by Allen and Shapanka.
- The estimation of these vectors required data on individual payouts and total payouts to contractors by project. Reports containing the data were available from FD's CPPS, which generates contractor reports from the names of contractors.[15] A random sample of approximately 30 projects was drawn per size group. There were less than 30 projects in three cases: contract group (7) had 10 projects, (8) had 4, and group (9) contained 29 projects in the sample.
- The contractor information was entered into an Excel spreadsheet. The first, second, and final payments were divided by the total payment to the contractor for each project to obtain the percentage of the contract cost that is paid out by these points in time. The resulting fractions were averaged over the sample to determine the vector entries.

ZD Vector (Contract Duration)

- This vector contains the expected duration (in months) for contracts in each of the nine contract size groups. Generally, contracts with relatively higher costs have longer duration.

---

[13] The number of payments ranged from a low of 215 in February to a high of 349 in January and had a mean of 277.

[14] The ratios of sample to population according to contract size were: (1)1.7%, (2) 4.2%, (3) 8.2%, (4) 16.9%, (5) 9.3%, (6) 16.9%, (7) 26.5%, (8) 17.7%, (9) 27.0%.

[15] Since FD did not have an historical list of contractors, the researcher opened a composite file of CPPS and printed this list. The contractor names were entered into the software and the corresponding reports were printed. The researcher tried to use only one project per contractor and get thirty projects per contract size group, to get an unbiased sample. This process was repeated until there were thirty samples in the majority of the contract size groups.

- Allen and Shapanka estimated a regression model that included original contract amount, month in which the contract was signed, road system, and project type as explanatory variables. Due to the difficulty of obtaining a like data set,[16] a different method was used in the update.
- The data came in hard copy from the Construction Division and included the size of the contract, final dollar value, and the completion date by project number and district from January 1, 1988 through December 9, 1994.
- The researcher searched the CD printouts for projects to fill the nine contract size groups. To compile a stratified random sample, the data set included four projects per district with completion years close to '88, '90, '92, and '94.[17]
- The vector entries are the average duration of projects within a particular contract size group, rounded up to the next highest integer. Therefore this calculation included only the size of the contract as an explanatory variable.

ZC Vector (Payout by Completion Date)

- The nine entries in this vector are employed to predict the amount paid out to projects by their completion date.
- The update of the ZC Vector was based on the original method used by Allen and Shapanka.
- The estimation of the vector required data on payments to completed projects and completion dates, both of which were available in the CPPS printouts from FD and the list of completed projects from CD. The CD printouts showed the completion dates by project number, and the CPPS printouts provided the payment schedules and total costs by project number. These two sets of data were examined to find matching project numbers and the data were entered into a spreadsheet file to have completion dates and payment schedules by project number in the same file.[18]
- The amount paid out by completion date was divided by the actual total payouts for each project. The ZC entries were computed by averaging the sample ratios in each of the contract size groups.

Consultations with FPDM indicated that the updated Z-Vectors did not produce results that followed the general pattern of the past several years. Since the updated vectors relied on data from FY 89 through FY 95 only, weighted averages of the pre-FY 89 vectors and the newly calculated vectors were obtained for each vector. Several

[16] Following this methodology for the current study would have been very difficult since hard copies of individual contracts are in storage and the authors were unable to discover their exact location. Alternatively, the authors could have relied on databases from ISD, but this method would have been very labor intensive since it would have involved combing through two data files of 3,823 and 6,463 records apiece. Furthermore, the data set still would have been incomplete since the month in which the contract was signed is not included in these ISD files.

[17] See Appendix B.

[18] See Appendix B.

different weights were calculated, then applied to each pair of vectors. These results were forwarded to FPDM in early April 1996.

Forecasting

Using the new Z Vectors for forecasting construction payout is relatively simple. The seven updated Z Vectors can be entered into the BASIC program in place of the older Z Vectors. Then the program may be run as before to provide forecasts of contract amounts paid out by VDOT each month.

Maintenance Expenditures Submodel

The maintenance expenditures submodel is used to forecast VDOT's payouts on maintenance each month. FPDM currently relies on the historical averages developed by Allen and Shapanka to predict these expenditures. Since payouts exhibited a relatively seasonal pattern, Allen and Shapanka developed a system of monthly factors to estimate maintenance.[19] The researchers have taken a new approach to develop the maintenance disbursements forecasts. This new approach relies on a regression model which was adapted for estimation purposes.

Data Requirements

The new methodology required the construction of a regression model that could be used to forecast maintenance expenditures accurately. The actual values of monthly maintenance expenditures, i.e., the dependent variable, came from FPDM and encompassed the period from FY 89 through FY 95 in current dollars.[20] The monthly maintenance disbursements follow a seasonal pattern of peaks in the summer and autumn months and troughs in the spring and winter months. The amount of money VDOT spends on maintenance may be influenced by many factors: the availability of money, the opportunity to spend it, and the conditions of the roadways. These factors can be categorized as: traffic characteristics, which affect the condition of roadways; economic factors, which are proxies for scarce traffic data; weather statistics, which affect both the condition of the roads and the opportunity to spend money; administrative features, which influence the availability of money and the rate at which it can be disbursed; and seasonal dummies, which proxy for weather or administrative factors.[21]

---

[19] Allen and Shapanka, p. 21.

[20] All money variables are in current dollars.

[21] The variables included in the forecasting equations are in Table 1, on page 14.

Traffic

All other things being equal, heavy traffic eventually leads to pavement deterioration. One would therefore expect an increase in traffic volume to be followed, after a certain period of time, by an increase in maintenance disbursements.

The following direct and indirect traffic variables were tested but were found to be insignificant:

- *Traffic Counts from Continuous Count Sites.* VDOT's Traffic Engineering Division (TED) calculates VMT on a monthly basis using assumptions based on historical findings. However, many or most of the continuous count sites are only a few years old. The data, even for just the 1990's, are stored in two or three databases and would be laborious to retrieve.[22] A time series may be available in the future; however, TED has no plans to store more than five or six years of continuous count data because of the prohibitive cost associated with the storage of such detailed data.
- *Net Taxable Motor Fuel Sales.* Available by month, and includes the amount of gas and diesel sold destined for highway use. Lagged values of this variable were tested.
- *A 65 mph Truck Speed Dummy.* As of July 1, 1994 the non-urban interstate speed limit for trucks is 65 mph. Trucks traveling at higher speeds may adversely affect the condition of the roadways. A variable capturing this change and perhaps providing information about maintenance spending would be an artificial variable, known in statistical terms as a "dummy variable," with values of one starting in July of 1994 and values of zero before that month.

Economic Factors

Since available direct measures of traffic volume are scarce, indirect measures of traffic volume may have some predictive value. For example, the Virginia Employment Commission (VEC) publishes a list of economic indicators each quarter, any of which may be correlated with, and thus serve as a proxy for, traffic volume.[23] The state of the economy affects choices made by consumers as well. Virginians may cut down on travel and use their vehicles less often during recessions. Interestingly, in periods of economic expansion, rising prices at the pump may discourage travel. These fluctuations in traffic and travel may affect the condition of roadways. Therefore, a measure of economic activity may be an important indicator of maintenance expenditures.

The following economic variables were tested but were found to be insignificant or unavailable:

---

[22] James B. Robinson of the Traffic Engineering Division provided the information on traffic counts.

[23] Virginia Economic Indicators. Virginia Employment Commission, Economic Information Services Division.

- *Gross State Product.* The Weldon Cooper Center for Public Service provided GSP. However, the series is only available by quarter, so an alternate economic variable had to be tested.
- *Taxable Retail Sales.* Available by month, this provides a picture of economic activity.[24] This series is one of the VEC's business indicators. Three, six, and twelve month lags of the variable were tested.
- *Taxable New Vehicle Registrations.* Available by month and a very strong indicator of business activity. Again, quarterly lagged values were tested.
- *Recession Dummy.* A tool used to determine if the recession of the early 1990's influenced expenditures on maintenance. The dummy variable will indicate in which cell or month an unusual event occurs. According to quarterly estimates of gross domestic product from the United States Statistical Abstracts, a recession occurred in the fourth quarter of calendar year 1990, and the first and second quarters of 1991.[25] For the sake of analysis then, the months of October, November, and December of 1990, and January, February, March, April, May, and June of 1991 were assigned a value of one.
- *Manufacturing Employment.* Available by month as an economic indicator.
- *Total Unemployment Rate.* Available on a monthly basis, and a VEC economic indicator.

The following economic variables were significant, i.e., they contributed to the explanation of maintenance disbursement behavior:

- *Nonagricultural Wage and Salary Employment.* Available by month and a VEC economic indicator.
- *Prime Rate.* Available by month from the U.S. Council of Economic Advisors' Economic Indicators.[26] Lagged values of the prime rate were also considered.

Weather

All other things being equal, certain kinds of weather such as low temperatures and heavy precipitation will mean some degree of pavement deterioration. One would expect these weather patterns to be followed, to some degree, by an increase in maintenance disbursements. Winter weather including snow and ice demands immediate maintenance response, but this very short response time makes it difficult to use data on these types of variables to predict maintenance expenditure ahead of time. Winter conditions may postpone routine maintenance activities. The following variables represent causative weather more or less directly:

---

[24] Taxable Retail Sales was provided by John Phisterer of the Virginia Department of Taxation.

[25] Statistical Abstract of the United States. 114th edition. 1994. The National Data Book,. United States Department of Commerce, Economics and Statistics Administration, Bureau of the Census.

[26] Economic Indicators. Prepared for the Joint Economic Committee by the Council of Economic Advisors. United States Government Printing Office, Washington.

- *Total Precipitation Per Month (water equivalent).* Available from National Oceanic and Atmospheric Administration publications.[27] One, two, three, four, five, and six month lags were also considered. Three and four month lags of precipitation were significant in only one of the forecasting equation alternatives.
- *Average Temperature by Month.* Also available from NOAA publications. One, two, three, four, five, and six month lags were also tested. Average temperature and a one month lag were only significant in one of the maintenance equation alternatives.

Administrative Factors

Established administrative practice may lead to predictable swings in maintenance disbursements which can be modeled by one of the seasonal factors mentioned above or even by another variable that follows the same seasonal pattern. Irregular, or regular but unpredictable administrative decisions will also affect maintenance expenditures.

The following administrative variables were considered but were found to be insignificant:

- *Annual Budget.* Expenditures by the Maintenance Division and the District Offices are guided to some degree by annual allocations. The annual budget for each year divided by twelve is a trend variable used to account for the tendency of the expenditures to rise from year to year due to inflation, etc. (This series was significant in conjunction with trigonomietric functions explained in *Seasonality Factors* below.)
- *Past (Lagged) Maintenance Expenditures By Month.* Future maintenance payouts may depend on payouts made in the past.
- *Carryforward Amounts.* Maintenance Division is permitted by law to carry over budget funds to cover maintenance replacement contracts entered in one year that will become payable in the following year. The size of the carryover presumably conveys information about the level of expenditure in the months of the following year. However, these figures are only readily available for FY 93 through FY 95.[28]
- *The Number of VDOT Employees.* Workforce totals by month were available from Management Services Division (MSD).

However, the following dummy variable was found to be significant during testing:

- *Biennium Cyclical Dummy.* A dummy variable that may be associated with the biennium budget cycle.

---

[27] Local Climatological Data: Monthly Summary. National Oceanic and Atmospheric Administration.

[28] Maintenance Carryforward amounts were tested for FY 93 through FY 95, and included in a separate set of step regressions along with the other variables for these years. The Carryforward amounts had insignificant t-statistics and were therefore not included in the model.

Seasonality Factors

Seasonal factors that reflect the average impact of weather at a given time of year may account for much of the actual movement in month-to-month maintenance expenditures. Maintenance disbursements follow a seasonal pattern, typically high in the summer and autumn months and relatively low in the winter and spring months. As the seasonal factors are based on historical trends rather than actual current weather data, they are presently available for forecasting maintenance months or years into the future. Seasonal variables may also proxy for some of the administrative practices that have a cyclical impact of maintenance payouts (see below).

The following variables were found to be insignificant:

- *Snow/Winter Dummies.* These allow a group of months when snow occurs to have its own weight. The snow dummy was assigned a value of one in months in which snowfall usually occurs and a value of zero in the remaining months. The "snow months" were varied to ensure the influencing months were captured by the dummy.
- *Trend Variable.* Allows for steady growth in the average level of expenditure.

The following variables were proven to be significant:

- *Month Dummies.* Similar to the ZM factors in the monthly factors submodel, i.e., they allow each month to have a different "weight" based on the historical trend.
- *Trigonometric Functions.* A sine wave that oscillates from high to low and back over the course of the year. It allows for the cyclical pattern of maintenance disbursements. One, two, and three month lags of these function values were also tested.

Modeling

The maintenance expenditure forecasts within FPDM have been based on historical patterns and monthly factors developed by FPDM staff. These factors are constantly evolving based on actual maintenance data. The forecasts that the Maintenance Division (MD) occasionally assembles are based on past monthly averages, modified by specific circumstances of the current year.[29]

The researchers tested two principal alternate devices to account for the seasonal pattern of maintenance disbursements. The first involved employing eleven month dummies, each of which would indicate the number of months remaining in the fiscal year. The second alternative was to use a trigonometric function. Both alternatives are discussed in the following sections. Besides these seasonal variables, intended to capture

[29] The Maintenance Division relies on historical patterns which are based on five years of data and adjusted by the experience of MD personnel to produce predictions of maintenance expenditure. (Personal communication, Gary Pokrifka and Lynn Poole, Maintenance Division.)

the year-in, year-out historical pattern in the maintenance figures, the researchers tested a variety of specific variables, described in detail in the previous section, that seemed likely to have value in predicting the future movement of maintenance expenditures.

Outliers

A few of the data points in the maintenance disbursement series appeared to be unusually low when compared to the same months in other years. For example, the amount for June 1989 was only $17,000, while for every other June the amounts ranged from $39,687,000 to $76,169,000. A similar phenomenon occurred in December of 1988. Conversations with Fiscal Division staff revealed that the reason for these anomalous entries was that the Commonwealth Accounting Report Systems (CARS) received flawed input into the system which generated a report for FY89 that contained erroneous information. To alleviate the skewing that these odd points caused, "breakpoint dummies" were used. These dummy variables have a value of one to correspond to the particular point in question, and a value of zero at every other point. In total, nine breakpoint dummies were tested but only the two named here contributed significantly to the maintenance disbursement models.

Month Dummy Models

These models incorporated the eleven month dummies and all the other variables in significance testing to build the appropriate regression model. This process involved running many preliminary regressions to determine which variables contributed significantly to the explanation of the behavior of maintenance disbursements. The dependent variable was used in a regression with each of the variables in turn. The variable with the highest t-statistic was chosen and then the dependent variable was run on this variable and all others, again in turn. This process continued until all the significant variables were chosen. Shortened variable nomenclature and the corresponding detailed variable title are shown in Table 1 below. Combinations of variables were also tested to determine if adding two or three variables at once could add to their significance. This process was repeated for every model formulation. The different models could then be compared and judged on the criterion of R^2. The model with the highest R^2 was:

$$MD = -105444 + 314.65Temp_{t-1} + 40.45NonAgEmp - 45714.1DJUNE89 - 21770DDEC88 + 2146.99JANUARY - 2061.86PrimeRate_{t-18} + 236.86Precip_{t-4} + 281.86Temp + 3714.65BiennDummy + 214.85Precip_{t-3} + \varepsilon \qquad (1)$$

where ε is a random disturbance. The disturbance arises for several reasons, but mainly because we cannot capture every influence on disbursements, such as all the inherent unpredictable elements in human behavior. The net effect of such intangible factors is captured in the disturbance term.

Table 1. Abbreviations for the Variables Used in Alternative Maintenance Models

NonAgEmp	*Nonagricultural Wage and Salary Employment*
DJUNE89	*Breakpoint Dummy for June 1989*
DDEC88	*Breakpoint Dummy for December 1988*
JANUARY	*Month Dummy for January*
$PrimeRate_{t-18}$	*Prime Interest Rate lagged 18 months*
$Precip_{t-4}$	*Total Precipitation lagged 4 months*
$Precip_{t-3}$	*Total Precipitation lagged 3 months*
Temp	*Average Temperature*
$Temp_{t-1}$	*Average Temperature lagged 1 month*
BiennDummy	*Biennium Cyclical Dummy*
AB	*Annual maintenance budget divided by 12*

The R^2 of the above model is 0.7147; in other words, equation (1) explains about 71% of the variation in maintenance disbursements. The regression model will have a slightly different form for forecasting purposes. This is similar to the trigonometric model case, as explained in the section on forecasting.

Trigonometric Function Models

The trigonometric function's purpose is to model the seasonal pattern of maintenance expenditures. It was conjectured that a sine wave with a period of one year and a phase shift of about pi/4 would imitate much of the seasonal variation. Such a sine wave relative to a July 1993 zero for FY 94 and 95 compared to maintenance disbursements is shown in Figure 1.

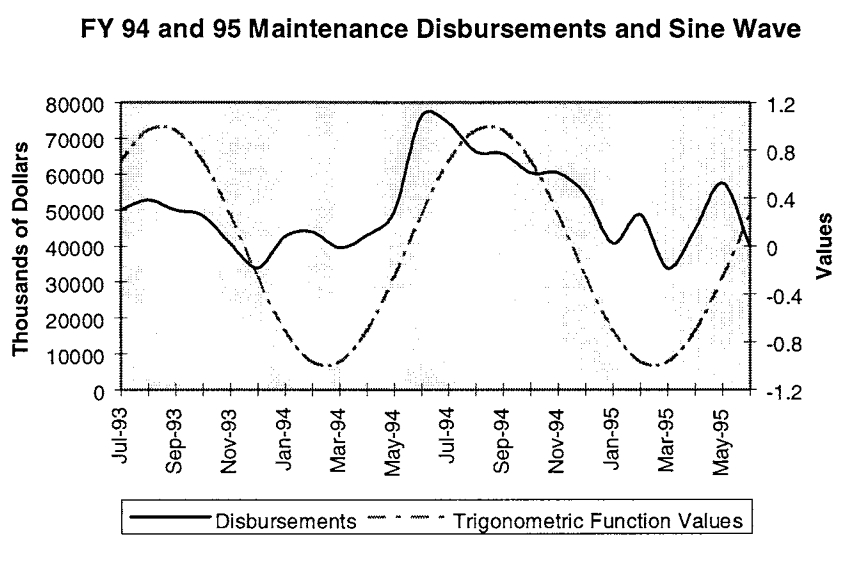


Figure 1.

The addition of a second term with a half year period and the same phasing adds more flexibility at the cost of estimating one additional parameter. The continuous trigonometric function is rendered into twelve discrete month-long pieces by integration. This means the area under the curve can be found. The equation for a typical month is:

$$y_m = \int_{\frac{\pi}{6}m-1}^{\frac{\pi}{6}m} [\sin(t+\theta) + A\sin(t+\theta)]dt \tag{2}$$

where
m=month index=1 for July, 12 for June, etc.
θ=Phase shift variable (different θ's move the sine wave laterally through time)
A=weight on the second term.

The value of the equation varied by month. For instance, the equation for July is:

$$y_{july} = \int_0^{\frac{\pi}{6}} [\sin(t+\theta) + A\sin(t+\theta)]dt \tag{3}$$

Once equation (3) was integrated, the form for our sample month was:

$$y_{july} = -\cos(\frac{\pi}{6}+\theta) - A/2\cos 2(\frac{\pi}{6}+\theta) + \cos(0+\theta) + A/2\cos 2(0+\theta) \tag{4}$$

To avoid the arduous task of iterating over both A and θ to determine which values of each result in the highest R^2 for the trigonometric function, the integrated equation was split into two parts:

$$y_{1july} = -\cos(\frac{\pi}{6}+\theta) + \cos(0+\theta) \tag{4a}$$

$$y_{2july} = -\cos 2(\frac{\pi}{6}+\theta) + \cos 2(0+\theta) \tag{4b}$$

It was considered possible that the size of the seasonal fluctuation might depend, to some extent, on the size of the annual budget. Various combinations of the trigonometric functions and annual budget figures were tested. The best form involved multiplying y1 and y2 by the annual budget figure divided by twelve:

$$MD = \alpha + \beta_1 y_1 AB + \beta_2 y_2 AB + \varepsilon \quad (5)$$

Because θ does not enter the equation linearly, the equation cannot be estimated by Ordinary Least Squares (OLS); the researchers iterated over different values of θ and then estimated the rest of the parameters by OLS.[30] Optimal values of θ were found for each form of the trigonometric equations. The trigonometric functions became columns of values once values for t were substituted into the equations.

These trigonometric equations were then put in the general regression and all variables were tested for significance. The following equation had the highest R^2 of all forms tested:

$$MD = -60585.1 - 46076.3_1 DJUNE89 - 30276.5DDEC88 + 0.0005y_1 AB + 0.00008y_2 AB$$
$$+40.2NonAgEmp - 1924.97\,PrimeRate_{t-18} + 3662.4BiennDummy + \varepsilon \quad (6)$$

The R^2 for this equation was 0.7076, therefore equation (6) explains about 71% of the variation in maintenance disbursements. This model has about the same R^2 as the month dummy model.

Since the two models have close values of R^2, they were judged on their forecasting accuracy, which is explained below. Both models were tested for heteroskedasticity (non-equal error variances) and autocorrelation (interdependent errors). The explanatory variables were tested for multicollinearity.[31] Tests for heteroskedasticity, autocorrelation, and multicollinearity revealed that both models are free of heteroskedasticity and autocorrelation. Equation (6) is free of multicollinearity, but equation (1) displays multicollinearity between two of its explanatory variables.[32]

The models were also tested for "structural break" by performing the Chow Test. Structural break may occur in a model if there was some element that fundamentally changed the data series, such as a war or a drastic change in policy. Both equation (1) and equation (6) were tested for structural break in the years FY 91 and FY 92. The Chow Test revealed that there is no structural break in either model.[33]

---

[30] By running repeated regressions with different values of θ the best value for θ can be reached. The corresponding values of A were found by dividing the coefficient for y2 by the coefficient for y1.

[31] The presence of any of these phenomena can lead to imprecise models and, therefore, imprecise and inaccurate forecasts. If some of these phenomena were present, one could increase the precision of the statistical analysis by correcting the estimation technique to explicitly account for such problems.

[32] A detailed discussion of the problems of heteroskedasticity, autocorrelation, and multicollinearity, as well as the particulars of their tests, appears in Appendix C.

[33] The details of the Chow Test are in the Appendix C.

Forecasting

Forecasting models differ slightly from the regression model on which they are based. Since each potential forecasting model contains at least one contemporaneous variable, for example nonagricultural wage and salary employment, this variable must be forecast before the model can be used for maintenance expenditure forecasts. Future values of the series may be calculated using average monthly growth rates. Thus, the alternate models for forecasting are as follows:[34]

$$MD = -60585.1 + 0.0005y_1 + 40.21(nonagemp_t * averagegrowthrate_t + nonagemp_t)$$
$$+0.00008y_2 - 1924.97\,\Pr ime_{t-18} + 3662.4BiennDummy \tag{7}$$

$$MD = -105444 + 314.65(templag1_t * avggrowthrate_t + templag1_t) +$$
$$40.45(nonagemp_t * avggrowthrate_t + nonagemp_t) + 2146.99JANUARY -$$
$$2061.86\,\Pr imeRate_{t-18} + 236.86\,\Pr ecip_{t-4} + 281.56(temp_t * avggrowthrate_t + temp_t)$$
$$+3714.65BiennDummy + 214.85\,\Pr ecip_{t-3} + \varepsilon \tag{8}$$

Forecasts for the period FY 90 through FY 95 were generated by substituting the relevant data points in each of the forecasting equations.[35] This procedure created two alternative sets of maintenance disbursements forecasts. After forecasts were generated using these models, the forecast errors and percentage differences for each of the two updated models were calculated to measure the accuracy of the models.[36] These errors were then compared to the errors produced by the current model for evaluation purposes. The errors were calculated over six years: FY 90 through FY 95. The forecast errors of each of the three models are shown below in Table 2. The percentage errors are shown in Appendix D. These forecast errors are calculated as the difference between the forecast and the actual value for each month. Then these figures are averaged for each month, which yields an average forecast error by month. This indicates the degree to which the forecasts have been "off" over the six years in the sample period. For example, for May, the current model used by FPDM yields an average forecast error of -$3,539,000 which means that the forecast has been underestimating the actual expenditures by this amount. In contrast, equation (7) underestimates for this month as well, but by $792,000. The absolute average forecast errors are averages of the absolute value of the monthly averages. They give an overall picture of the performance of each model.

---

[34] Since the two breakpoint dummies are used to account for anomalous events, they are not included in the forecasting equation.

[35] FY 89 was excluded from the forecast period because the researchers did not have access to the necessary data for inclusion in the models prior to FY 89. In addition, the reported actual maintenance expenditures for FY 89 are inaccurate for December of 1988 and June of 1989.

[36] A template for generating maintenance expenditure forecasts is in Appendix E.

Table 2. Comparison of Model Errors

	Month	Average Monthly Forecast Error (Thousands of Dollars)
	July	-3,432
	August	1,859
	September	4,731
	October	9,189
	November	-4,540
Current Model	December	-3,680
	January	5,558
	February	5,271
	March	1,188
	April	-909
	May	-3,539
	June	1,388
	Average Absolute Error	*3,744*
	July	-36471
	August	-788
	September	3,522
	October	-1,522
	November	-733
Best Month Dummy Model	December	-3,550
	January	2,780
	February	508
	March	2,474
	April	3,337
	May	-941
	June	-2,2615
	Average Absolute Error	*2,172*
	July	-2,357
	August	208
	September	1,890
	October	-3,488
	November	-28
Best Trigonometric Model	December	2,710
	January	-1,259
	February	2,492
	March	-303
	April	-1,048
	May	-792
	June	134
	Average Absolute Error	1,392

Since the trigonometric model (equation 7) produced smaller overall forecast errors, is simpler to use than equation (8), and is free of multicollinearity, it was chosen by the researchers as the appropriate forecasting model.

FINDINGS AND RECOMMENDATIONS

A re-estimation of the Monthly Factors Submodel in a similar fashion to the original study yielded Z-Vector factors similar, but not identical to the values currently used by FPDM (last updated in 1988). Since the confidence interval for the 1988 estimates was not published, it is not possible to state how many of the new estimates differ significantly from the old. The budget line item for construction that appears in

FPDM reports is a combination of separate totals, namely construction contracts, state force construction, and consultant contracts and miscellaneous contracts, each forecast by a different method. Although the actual amount spent on each construction contract is recorded, and a sample of the contracts was used to re-estimate some of the parameters of the Monthly Factors Submodel, obtaining total construction contract payouts per month separate from the other totals in the construction line item proved impractical. These facts make it impossible to compare directly the forecasting performance of the current FPDM submodel with that of the updated submodel. It is possible that the current and updated Z-Vectors represent two separate estimates of a relationship that has not changed very much; if this is the case, to abandon the current Z-Vector estimates would be a waste of valuable information.

1) Since the historical statistical evidence on actual construction contract payouts is difficult to retrieve, it is recommended that FPDM continue to generate contract payouts by the existing submodel, and begin to generate a competing set of forecasts by a submodel incorporating a 50:50 weighted average of the current Z-Vectors and the re-estimated ones. This practice will generate over time a series of parallel "Highway Systems Acquisition and Construction" forecasts whose performance can be compared directly. The historical data may become more accessible as VDOT's computer databases become increasingly advanced.

2) The estimation of a new maintenance expenditures equation combining seasonal factors based on historical averages similar to those in use currently, with leading indicators not previously included, produced a new submodel that performs markedly better than the submodel currently in use. In the FY 90 through 95 period, forecasts by the existing method differed from the actual monthly maintenance outlays by an average of $3.774 million. Over the same period, forecasts by the best revised model differed from the monthly actuals by an average of $1.392 million -- about 37% of the error under the existing method. Given these findings, it is recommended that FPDM adopt the submodel represented as equation (7) for future maintenance expenditures forecasts. VTRC will provide FPDM with an Excel spreadsheet that has all the components of the maintenance forecasting model embedded in its cells.

3) Finally, to continuously refine and adapt the model to the continuing effects of both internal and external forces, it is recommended that VTRC, in cooperation with FPDM, continue to monitor the performance of the newly updated submodels. It is also recommended that the components of the Cash Flow Forecasting Model be reviewed again in about three years.

REFERENCES

1. Allen, G. R., and Shapanka, A. (1983) *A System for Monitoring Cash Flow: Phase I: Forecasting Payments on Construction Contracts*. Charlottesville: Virginia Highway and Transportation Research Council.

2. Allen, G. R., and Shapanka, A. (1985) *A System for Monitoring Cash Flow: Phase II: Forecasting Federal and State Revenues, Contracts, Other Expenditures, and Cash Balances.* Charlottesville: Virginia Highway and Transportation Research Council.

3. *Virginia Economic Indicators*. Volumes 20 through 28. Virginia Employment Commission: Economic Services Division. Richmond, Virginia.

4. *Local Climatological Data: Monthly Summary*. (1987 through 1995) National Oceanic and Atmospheric Administration, National Environmental Satellite, Data and Information Service. Asheville, North Carolina: National Climatic Data Center.

5. *Statistical Abstracts of the United States 1994. The National Data Book. 114th edition*. (1994) United States Department of Commerce: Economics and Statistics Administration. Washington, D.C.: Bureau of the Census.

6. *Economic Indicators*. (1986 through 1995) Prepared for the Joint Economic Committee by the Council of Economic Advisors. Washington, D.C.: United States Government Printing Office.

7. Greene, W. H. *Econometric Analysis*. (1990) New York: MacMillan Publishing Inc.

APPENDICES

APPENDIX A: Cash Flow Forecasting Model Product

APPENDIX B: Monthly Factors Submodel Development

APPENDIX C: Maintenance Expenditures Submodel: Model Development and Tests

APPENDIX D: Maintenance Expenditures Submodel: Model Performance

APPENDIX E: Maintenance Expenditures Submodel: Forecast Template

Appendix A

Cashflow Forecasting Model Product

Linking the Cash Flow Forecasting Model and its Submodels

Submodel	Cash Flow Line Item
State Revenue	Total State Sources
Federal Aid Reimbursement	Federal Aid
Monthly Factors Model	Highway Systems Acquisition and Construction
Construction Model	same as above
Consultant Contracts, Miscellaneous Contracts, and Other Expenditures	same
Maintenance Expenditures Model	Highway Systems Maintenance
Expenditures on Materials and Supplies	same as above
Payments to Localities, Other Agencies, and Transit Properties, and Expenditures on Salaries and Wages, Equipment, and Right-of-Way	Financial Assistance to Localities, Mass Transit, Support to Other State Agencies and Land Management

Notes:
1. Three of the submodels link up to the same cash line item. The results from the Monthly Factors Model, the Construction Model, and the Consultant Contracts, Misc. Contracts, and Other Expenditures Model are all part of the Highway Systems and Acquisition and Construction line item. Two other submodels link to the Highway Systems Maintenance line item.
2. In the Payments to Localities, etc., submodel, each component is estimated, and the result is then entered into the corresponding cash flow line item.
3. The information for this schematic was obtained from discussions with Ms. Diane Mosby of FPDM.

October '93 Forecast

Fund		Estimate JUL 1993	Estimate AUG 1993	Estimate SEP 1993	Estimate OCT 1993	Estimate NOV 1993	Estimate DEC 1993	Estimate JAN 1994	Estimate FEB 1994	Estimate MAR 1994	Estimate APR 1994	Estimate MAY 1994	Estimate JUN 1994	TOTAL FISCAL
HMOF	BEGINNING CASH & INVESTMENTS	219,919	174,882	180,356	179,113	134,300	139,546	149,158	121,264	124,999	145,252	132,778	140,314	219,919
HMOF														
HMOF	REVENUES													
HMOF	STATE SOURCES													
HMOF	Motor Vehicle Fuel Tax	28,789	49,281	44,572	51,552	44,402	45,729	47,785	44,162	42,157	41,497	47,436	60,104	547,466
HMOF	Motor Vehicle Sales & Use	15,438	15,995	15,262	14,734	13,013	11,540	14,221	12,819	16,744	17,181	16,769	18,263	181,978
HMOF	State Sales & Use Tax	0	0	0	0	0	0	0	0	0	0	0	0	0
HMOF	Motor Vehicle Licences	9,019	9,245	7,961	8,235	6,602	6,192	8,296	8,366	9,608	9,709	10,169	9,703	103,104
HMOF-	Interest (3.750%)													0
HMOF	Other	4,045	2,217	1,219	1,373	1,110	1,477	2,263	3,679	11,088	6,739	3,891	3,509	42,610
HMOF														
HMOF	TOTAL STATE SOURCES	57,290	76,738	69,014	75,895	65,126	64,937	72,564	69,027	79,596	75,125	78,265	91,579	875,158
HMOF														
HMOF	FEDERAL AID	795	795	795	795	795	795	795	795	795	795	795	795	9,536
HMOF	LOCAL CONTRIBUTION	0	0	0	0	0	0	0	0	0	0	0	0	0
HMOF	TOLL REVENUE	0	0	0	0	0	0	0	0	0	0	0	0	0
HMOF														
HMOF	TOTAL REVENUES	58,085	77,533	69,809	76,689	65,921	65,732	73,359	69,822	80,391	75,920	79,060	92,374	884,694
HMOF	TRANSFERS IN	0	0	0	0	0	0	0	0	0	0	0	0	0
HMOF														
HMOF	TOTAL REVENUES & TRANSFERS IN	58,085	77,533	69,809	76,689	65,921	65,732	73,359	69,822	80,391	75,920	79,060	92,374	884,694
HMOF														
HMOF	AVAILABLE FOR DISBURSEMENT	278,004	252,414	250,164	255,802	200,221	205,277	222,517	191,086	205,390	221,171	211,838	232,688	1,104,613
HMOF														
HMOF	EXPENDITURES (By Program)													
HMOF	Admin. and Support Services	2,863	4,295	7,158	7,158	5,727	3,579	5,727	7,158	6,442	5,011	6,442	10,022	71,583
HMOF	Ground Trans. Regulation	595	595	595	595	595	595	595	595	595	595	595	595	7,139
HMOF	Ground Trans. Sys. P & R	238	238	238	238	238	238	238	238	238	238	238	238	2,851
HMOF	Highway Sys. Acq. & Cons.	0	0	0	0	0	0	0	0	0	0	0	0	0
HMOF	Highway Sys. Maintenance	58,977	61,090	58,743	66,430	42,311	37,675	46,360	36,912	33,743	37,793	45,128	61,794	586,958
HMOF	Ground Trans. Sys. Safety	536	536	536	536	536	536	536	536	536	536	536	536	6,434
HMOF	Toll Facility Operations	0	0	0	0	0	0	0	0	0	0	0	0	0
HMOF	Fin. Assist. to Localities	35,921	0	0	42,763	0	1,711	41,052	0	1,711	39,342	1,711	5,132	169,340
HMOF	Mass Transit	0	0	0	0	0	0	0	0	0	0	0	0	0
HMOF	Support to Other State Ag.	1,042	1,042	1,042	1,042	1,042	1,042	1,042	1,042	1,042	1,042	1,042	1,042	12,509
HMOF	Land Management	50	50	50	50	50	50	50	50	50	50	50	50	598
HMOF	Capital Outlay	0	0	0	0	0	0	0	0	0	0	0	0	0
HMOF														
HMOF	TOTAL EXPENDITURES	100,222	67,846	68,362	118,812	50,499	45,426	95,600	46,531	44,357	84,606	55,742	79,408	857,412
HMOF	OTHER DISBURSEMENTS (Incl GLA)	(381)	1,142	(381)	(381)	1,142	(381)	1,485	1,142	(381)	(381)	(381)	(381)	1,866
HMOF	TRANSFERS OUT	3,281	3,071	3,071	3,071	9,034	11,074	4,168	18,413	16,162	4,168	16,162	19,166	110,841
HMOF														
HMOF	AMOUNT DISBURSED	103,123	72,059	71,052	121,502	60,675	56,119	101,253	66,086	60,139	88,393	71,524	98,194	970,119
HMOF														
HMOF	NET CASH FLOW	(45,038)	5,474	(1,243)	(44,813)	5,246	9,613	(27,894)	3,736	20,252	(12,474)	7,537	(5,820)	(85,425)
HMOF														
HMOF	ENDING CASH & INVESTMENTS	174,882	180,356	179,113	134,300	139,546	149,158	121,264	124,999	145,252	132,778	140,314	134,494	134,494
HMOF														

APPENDIX B

Monthly Factors Submodel
Submodel Development

APPENDIX B

MONTHLY FACTORS SUBMODEL

Equations used to estimate Z-Vectors

ZA Vector

$$ZA = 1/n\left(\sum_{i=1}^{9} FinalContractAmount_i \div \sum_{i=1}^{9} EstimatedContractAmount_i\right)$$,where

i=contract size group, i=1,2,3,4,5,6,7,8,9.

ZM Vector

$$ZM = 1/v\left(\sum_{\tau=-1}^{n} ConstructionContractPayout_\tau \div \sum_{\kappa=1}^{n} ConstructionContractPayout_\tau\right)$$

where 1/v is a weighted average, where τ = month and κ = fiscal year, κ = FY 87 through FY 94.

ZF, ZS, and ZL Vectors

$$ZF = 1/n\left(\sum_{\rho}^{n} FirstPayment_i \div TotalPayments_i\right)$$, where ρ= individual contracts

$$ZS = 1/n\left(\sum_{\rho}^{n} SecondPayment_i \div TotalPayments_i\right)$$

$$ZL = 1/n\left(\sum_{\rho}^{n} LastPayment_i \div TotalPayments_i\right)$$

ZC Vector

$ZC = 1/n\left(\sum AmountPaidOutbyCompletionDate_i \div TotalContractAmount_i\right.$ where i = contract size group

ZD Vector

$$ZD = 1/n\sum \left(FirstPaymentmonth_i - CompletionMonth_i\right)$$

Sample Sizes

ZA Vector Sample Sizes

The initial sample size and the sample size used after incomplete contracts, blank cells, and outliers were discarded are, respectively, by contract group: (1) 1810, 1580, (2) 713, 638, (3) 366, 314, (4) 177, 160, (5) 321, 285, (6) 177, 155, (7) 113, 101, (8) 34, 30, (9) 111, 95.

ZD Sample Size

The sample sizes for the nine groups are: (1) 36, (2) 36, (3) 34, (4) 35, (5) 34, (6) 28, (7) 19, (8) 5, (9) 10.

ZC Sample Size

The sample sizes for the nine contract groups are: (1) 10, (2) 12, (3) 11, (4) 10, (5) 10, (6) 13, (7) 7, (8) 5, (9) 10.

ZF, ZS, and ZL Vectors

All contract groups but three had 30 projects in the sample: (7) 10, (8) 4, and (9) 29.

How Z-Vectors are used in BASIC program

$ZA_\iota \times$ Initial Contract Amount = Final Contract Amount

$ZM_\iota \times$ Total Payouts for the Year = Payout for an Individual Month

$ZF_\iota \times$ Total Contract Amount = First Payout
$ZS_\iota \times$ Total Contract Amount = Second Payout
$ZL_\iota \times$ Total Contract Amount = Last Payout

$ZC_\iota \times$ Total Contract Amount = Amount to be Paid by Completion Date

$ZD_\iota f$(contract size) = Expected Duration of a Project

Z-Vectors

ZM Vector

	New Values	Current Values*
July	0.1048	0.0987
August	0.0961	0.0919
September	0.0999	0.1044
October	0.1038	0.1056
November	0.0908	0.0950
December	0.0987	0.0933
January	0.0751	0.0674
February	0.0408	0.0460
March	0.0621	0.0519
April	0.0703	0.0767
May	0.0760	0.0809
June	0.0862	0.0975

ZD Vector

Contract Size	New Values	Current Values
< $300,000	5	6
>$300,000 <$500,000	7.5	10
>$500,000 <$750,000	8	10
>$750,000 <$1,000,000	9	11
>$1,000,000 <$2,500,0	10	11
>$2,500,000 <$5,000,0	13	15
>$5,000,000 <$8,000,0	15	18
>$8,000,000 < $10,000,	17	20
>$10,000,000	22.5	24

ZC Vector

Contract Size	New Values	Current Values
<$300,000	0.6650	0.5500
>$300,000 <$500,000	0.6900	0.6900
>$500,000 <$750,000	0.8100	0.8100
>$750,000 <$1,000,000	0.8500	0.8500
>$1,000,000 <$2,500,000	0.8850	0.8900
>$2,500,000 <$5,000,000	0.9250	0.9300
>$5,000,000 <$8,000,000	0.9750	1.0000
>$8,000,000 <$10,000,000	0.9850	1.0000
>$10,000,000	0.9650	1.0000

ZA Vector

Contract Size	New Values	Current Values
<$300,000	1.0795	1.1400
>$300,000 <$500,000	1.0540	1.1100
>$500,000 <$750,000	1.0720	1.1700
>$750,000 <$1,000,000	1.0880	1.1800
>$1,000,000 <$2,500,0	1.0825	1.1500
>$2,500,000 <$5,000,0	1.0005	1.0300
>$5,000,000 <$8,000,0	0.9945	1.0200
>$8,000,000 <$10,000,	1.0285	1.0200
>$10,000,000	1.0340	1.0200

ZF, ZS, and ZL Vectors

	ZF	ZS	ZL	ZF	ZS	ZL
Contract Size	New Values	New Values	New Values	Current Values	Current Values	Current Values
< $300,000	0.0794	0.1558	0.1343	0.0240	0.0890	0.2080
>$300,000 <$500,000	0.0433	0.1127	0.0651	0.0100	0.0410	0.0920
>$500,000 <$750,000	0.0435	0.0991	0.0691	0.0100	0.0500	0.1000
>$750,000 <$1,000,000	0.0520	0.0730	0.0881	0.0100	0.0220	0.1200
>$1,000,000 <$2,500,000	0.0390	0.0665	0.0538	0.0100	0.0400	0.0700
>$2,500,000 <$5,000,000	0.0331	0.0527	0.0348	0.0100	0.0350	0.0580
>$5,000,000 <$8,000,000	0.0276	0.0378	0.0318	0.0100	0.0150	0.0400
>$8,000,000 < $10,000,000	0.0203	0.0259	0.0220	0.0100	0.0150	0.0290
>$10,000,000	0.0182	0.0343	0.0344	0.0100	0.0180	0.0340

*"Current Vectors" are the numbers used by FPDM at present.

APPENDIX C

Maintenance Expenditures Submodel
Model Development and Tests

Heteroskedasticity, Autocorrelation, and Multicollinearity

Heteroskedasticity occurs when the error terms have different variances. In the case of this project, this implies that the variance of maintenance disbursements changes over time. This means that the results of the regression will be biased, therefore the coefficients will be biased. The estimates of the coefficients will be either too low or too high and therefore the predictive quality of the model will suffer. For example, suppose a model with heteroskedasticity produced a coefficient value of 3 for the variable temperature. This coefficient would be inaccurate and therefore the forecast would be inaccurate. Autocorrelation occurs when the disturbance term relating to any observation is influenced by the disturbance term relating to any other observation. For example, if there is autocorrelation, a disruption in one of the variables may be carried over into the next month. The presence of autocorrelation will also result in biased estimators. Multicollinearity occurs when there is a linear relationship between some or all the explanatory variables. If perfect multicollinearity is present, the regression coefficients of the explanatory variables are indeterminate and their standard errors are infinite. If multicollinearity is less than perfect, the coefficients possess large standard errors, which means the coefficients cannot be estimated with great accuracy.

Tests for Heteroskedasticity

The problem of heteroskedasticity is more likely to appear in models with cross-sectional data than in models with time series data. Even though all the formations of the maintenance model are based on time series data, both forecasting models were tested for the presence of heteroskedasticity in two ways. First, the Glejser test was performed.[1] The hypothesis of the test is that is that the heteroskedasticity in the model comes from one of the independent variables. The absolute values of the residuals from each model are regressed against different functional forms of each of the explanatory variables and an intercept. A significant coefficient on any one of the explanatory variables indicates the presence of heteroskedasticity. The Glejser test revealed that none of the coefficients were significant - in fact, they were highly insignificant, which means the hypothesis is disproved and there is no heteroskedasticity in either model.

A second test that is more general in nature than the Glejser test, known as the White Test, was also performed on each model. The White test is an indicator of *any* heteroskedasticity between any variable or combination of variables in the model. The White test will also indicate any misspecification in the model.[2] The squared residuals from each model were regressed on as many variables and combinations of variables as Excel could accommodate at one time (17). The R^2 of this regression was multiplied by the number of observations (84). This critical value was then compared to the test value of the chi-square at a 95% level of confidence with 16 degrees of freedom. The value from the test on the trigonometric model was 12.47 and the critical chi-square value was 26.29. Since $12.47<26.29$, the hypothesis is rejected and therefore no heteroskedasticity was found in this model. For the month dummy model, the test value was 10.66, and since this is less than the critical value, the test finds no heteroskedasticity.

---

[1] Greene, William H. *Econometric Analysis*, page 423.

[2] Greene, page 419.

Test for Autocorrelation

The Durbin-Watson Test was performed on the model. The test is based on the calculation of the *d statistic*, which is the ratio of the sum of squared differences in successive residuals (errors) to the sum if the squared residuals. Since the d statistic is computed from the residuals which are dependent on the explanatory variables, the distribution of the *d* is difficult to find. Unlike other tests such as the t-test or the F-test, there is no unique critical value that will lead to the acceptance or rejection of the hypothesis that there is no autocorrelation. The *d* should not differ from the value of 2 - otherwise there is autocorrelation. The performance of the test on the trigonometric model resulted in a value of 2.13, which is not significantly different from 2, therefore, there is no autocorrelation this model. The Durbin-Watson test on the month dummy model resulted in a value of 2.09, which is not significantly different from 2, therefore there is no autocorrelation in this model either.

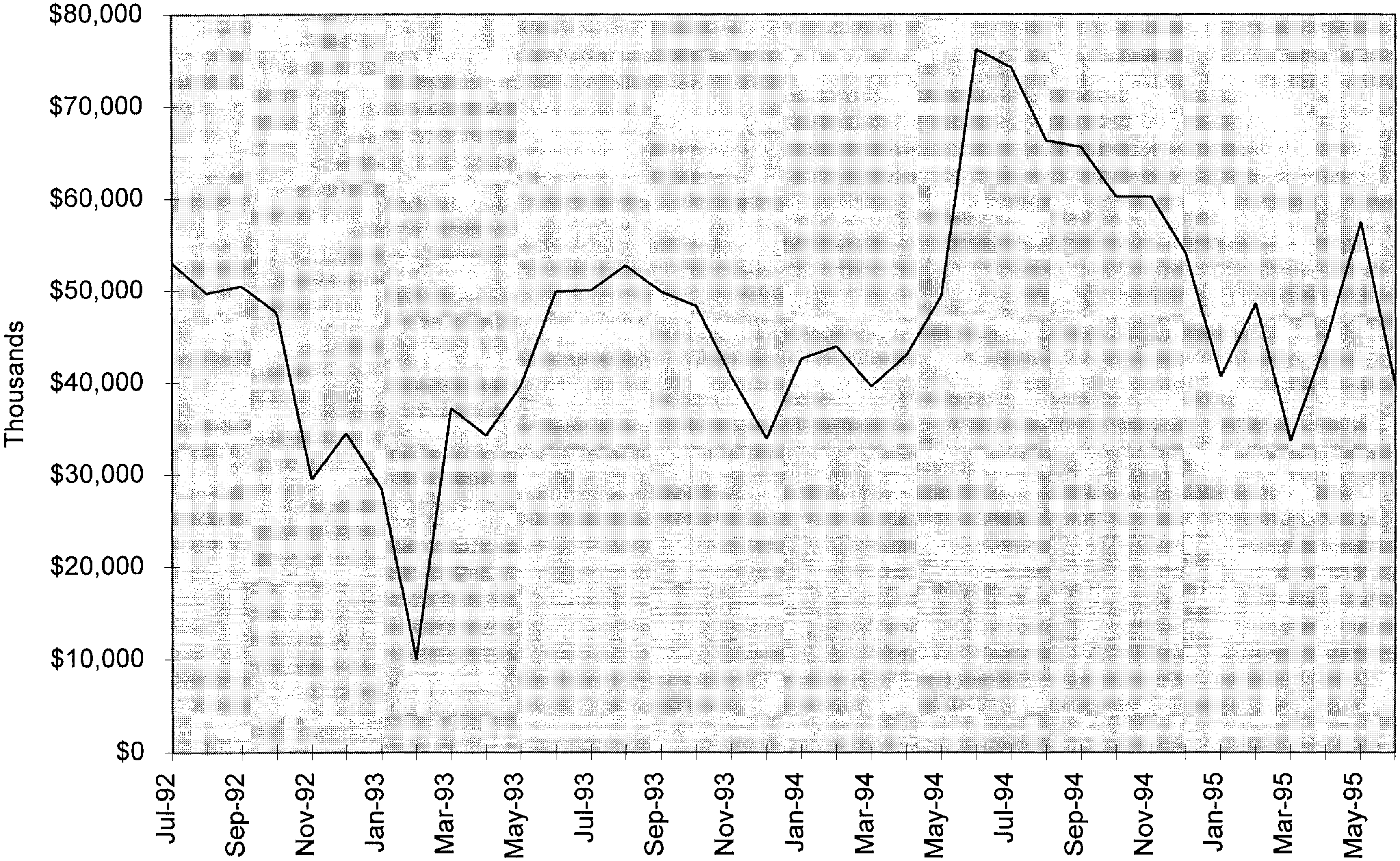
Test for Multicollinearity

Multicollinearity exists when the correlation coefficient between variables approaches a value of one. In practice, any value above 0.7 can cause problems in a regression model. In the case of the trigonometric model, the highest correlation coefficient for any combination was only 0.65. Therefore, there is an absence of multicollinearity in this particular model. However, the month dummy model has a correlation coefficient of 0.83 between the variables temperature and temperature lagged one month. This indicates some multicollinearity between these two variables. Accordingly, the coefficients may not be estimated with a great deal of accuracy.

Structural Break Test

The Chow Test is a method of determining the presence of structural break in a regression model. A structural break may occur due to a significant disruptive event or an exogenous shock such as severe weather, an international conflict, or a change in policy. To be sure that the organizational changes experienced by VDOT in the early 90's did not cause a break in the forecasting models, test were performed for breaks in FY 91 and FY 92. To perform the test, the sample was divided into two parts. A pooled regression is run with all the observations, and then two regressions are run using each of the two sample parts. The residual sum of squares of all three regressions are compared according to the test formula. The test statistic is based on a chi-square distribution with observation and parameter-specific degrees of freedom in the numerator and denominator. For the FY 91 break test on the month dummy model, the test value was 1.8. This is less than the chi-square critical value of approximately 2.04 at 9 and 66 degrees of freedom. The FY 92 break test value was 0.6, which is less than the critical value. This indicates an absence of structural break, i.e., for both years in question, the two regressions are statistically the same. For the trig model, the test value for FY 91 was 0.91, and for FY 92 was 0.74. Since both break tests are below the relevant critical value, there is no indication of a structural break in either of these years.

Maintenance Disbursements FY 93 Through FY 95



SUMMARY OUTPUT

Trigonometric Model

Regression Statistics	
Multiple R	0.8412
R Square	0.7076
Adjusted R Square	0.6807
Standard Error	8182.1721
Observations	84.0000

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	7.0000	12315466700.3381	1759352385.7626	26.2794	0.0000
Residual	76.0000	5088043454.5548	66947940.1915		
Total	83.0000	17403510154.8929			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-60585.1005	46322.8633	-1.3079	0.1949	-152845.1968	31674.9957	-152845.1968	31674.9957
DJU89	-46076.2522	8365.3878	-5.5080	0.0000	-62737.3872	-29415.1171	-62737.3872	-29415.1171
DD88	-30276.5235	8468.2104	-3.5753	0.0006	-47142.4477	-13410.5993	-47142.4477	-13410.5993
Trig 1	0.0005	0.0001	8.6063	0.0000	0.0004	0.0006	0.0004	0.0006
Non ag wage and Sal Empl	40.2054	14.5937	2.7550	0.0073	11.1395	69.2712	11.1395	69.2712
Trig 2	0.0001	0.0000	2.7858	0.0067	0.0000	0.0001	0.0000	0.0001
prime18	-1924.9732	687.2872	-2.8008	0.0065	-3293.8260	-556.1203	-3293.8260	-556.1203
Election Year Dummy	3662.3969	1829.4158	2.0019	0.0489	18.7948	7305.9990	18.7948	7305.9990

SUMMARY OUTPUT

Month Dummy Model

Regression Statistics	
Multiple R	0.8454
R Square	0.7147
Adjusted R Square	0.6757
Standard Error	8246.7053
Observations	84.0000

ANOVA

	df	SS	MS	F	Significance F
Regression	10.0000	12438915386.5379	1243891538.6538	18.2903	0.0000
Residual	73.0000	4964594768.3550	68008147.5117		
Total	83.0000	17403510154.8929			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-105443.8888	48058.4489	-2.1941	0.0314	-201224.3043	-9663.4732	-201224.3043	-9663.4732
Temp lag 1	314.6519	120.7950	2.6048	0.0111	73.9076	555.3962	73.9076	555.3962
Non ag wage and Sal Empl	40.4539	15.2310	2.6560	0.0097	10.0985	70.8092	10.0985	70.8092
DJU89	-45714.0515	8554.7435	-5.3437	0.0000	-62763.6424	-28664.4607	-62763.6424	-28664.4607
Dummy #7	2146.9931	730.0875	2.9407	0.0044	691.9299	3602.0563	691.9299	3602.0563
DD88	-21769.9757	8695.0190	-2.5037	0.0145	-39099.1354	-4440.8161	-39099.1354	-4440.8161
prime18	-2061.8638	704.9707	-2.9248	0.0046	-3466.8692	-656.8583	-3466.8692	-656.8583
precip lag 4	236.8559	105.2834	2.2497	0.0275	27.0263	446.6855	27.0263	446.6855
Election Year Dummy	3714.6537	1847.5627	2.0106	0.0481	32.4641	7396.8433	32.4641	7396.8433
Average Temperature	281.8580	121.7079	2.3159	0.0234	39.2943	524.4216	39.2943	524.4216
precip lag 3	214.8487	107.6731	1.9954	0.0497	0.2563	429.4411	0.2563	429.4411

Correlation Matrices for Trigonometric Model and Month Dummy Model Variables

Trigonometric Model

	Trig 1	Non ag wage and Sal Empl	Trig 2	prime18	ElectYear Dummy
Trig 1	1.0000				
Non ag wage and Sal Empl	0.0967	1.0000			
Trig 2	0.0000	0.0262	1.0000		
prime18	0.0075	-0.6543	-0.0030	1.0000	
Election Year Dummy	0.0915	0.0215	0.0000	0.0603	1.0000

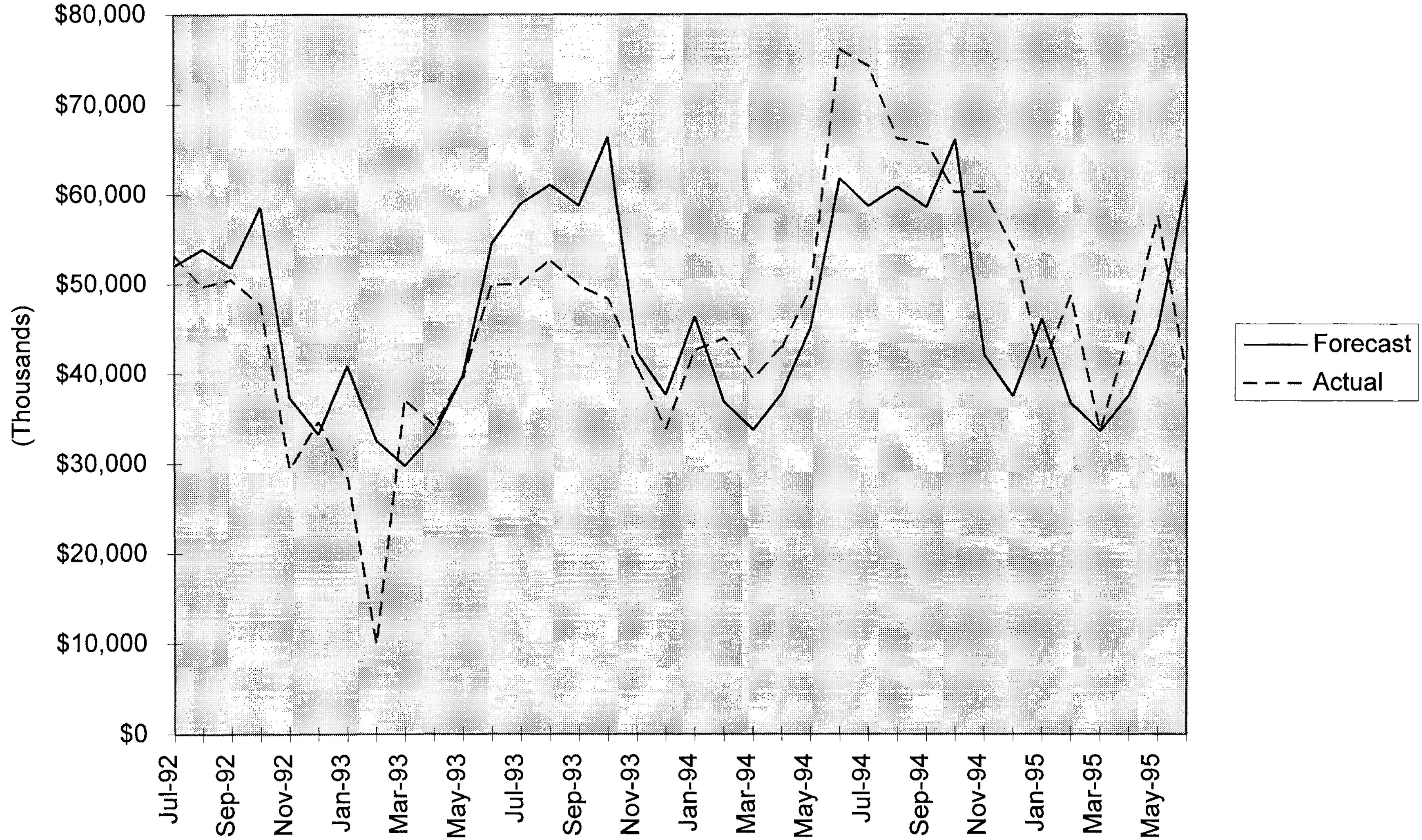
Month Dummy Model

	Temp lag 1	Non ag wage and Sal Empl	DJU89	Dummy #7	DD88	prime18	precip lag 4	Election Year Dummy	Average Temperature	precip lag 3
Temp lag 1	1.0000									
Non ag wage and Sal Empl	0.0960	1.0000								
DJU89	0.0362	0.0005	1.0000							
Dummy #7	-0.3856	-0.1662	-0.0331	1.0000						
DD88	-0.0688	-0.0822	-0.0120	-0.0331	1.0000					
prime18	0.0326	-0.6543	0.0279	-0.0003	-0.0426	1.0000				
precip lag 4	0.1460	0.1210	0.0106	-0.1099	0.0084	0.0122	1.0000			
Election Year Dummy	0.0649	0.0215	-0.1098	-0.0431	0.1098	0.0603	0.1157	1.0000		
Average Temperature	0.8319	0.0877	0.1323	-0.3930	-0.1492	0.0323	0.1519	0.0419	1.0000	
precip lag 3	0.2264	0.1010	0.0866	-0.1337	-0.0988	0.0238	0.0491	0.0219	0.1477	1.0000

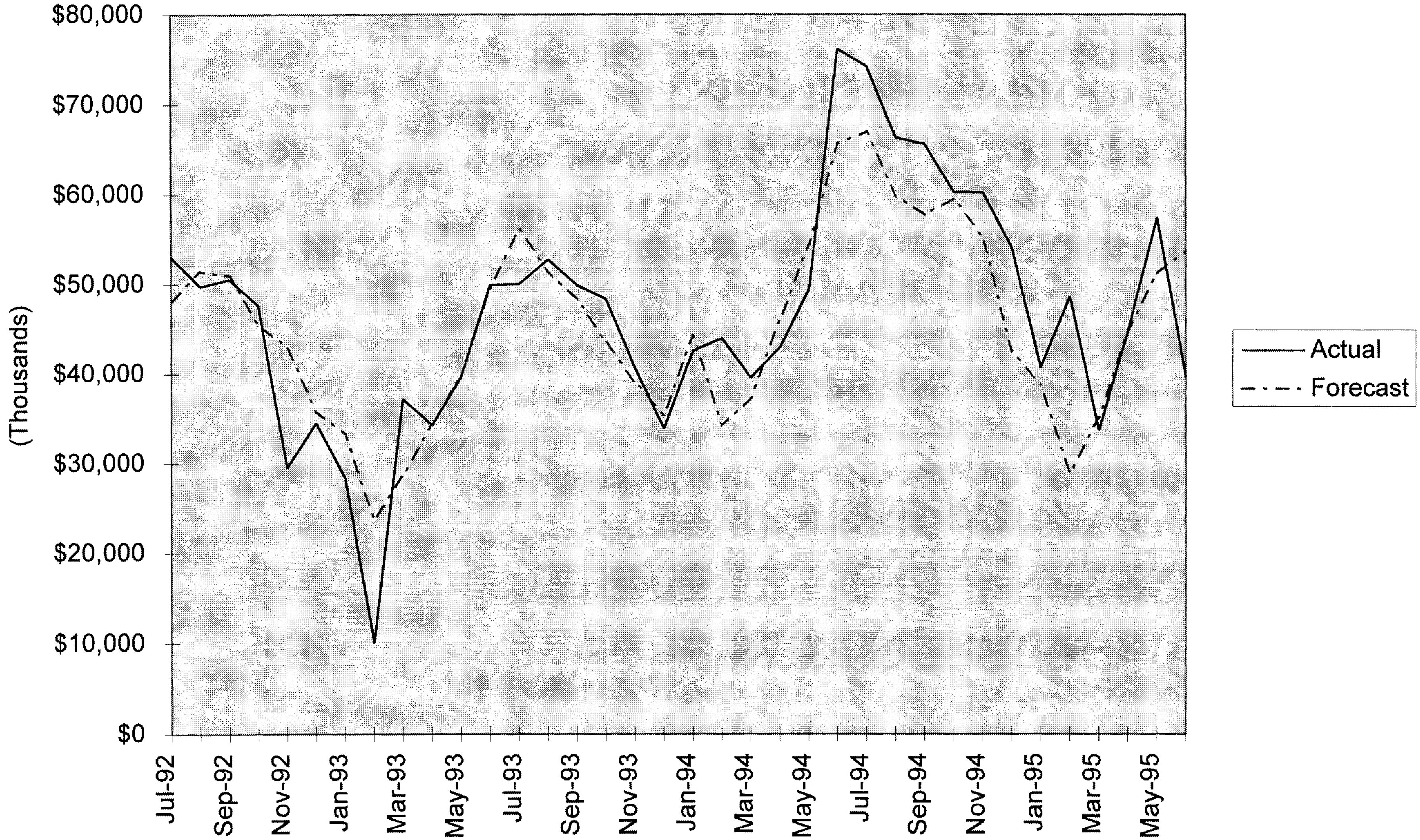
APPENDIX D

Maintenance Expenditures Submodel
Model Performance

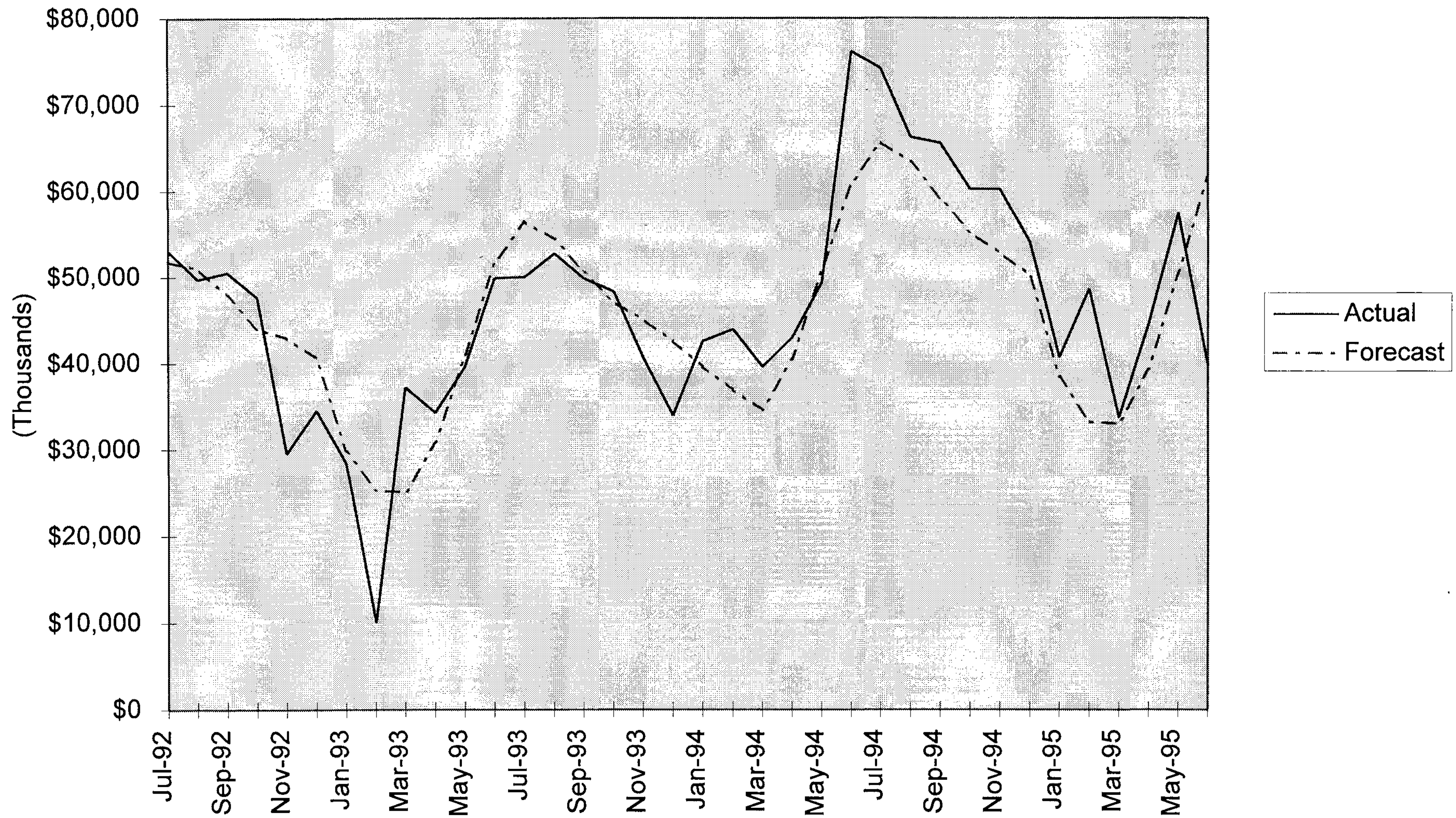
Forecast vs. Actuals Using Current Model



Forecast vs. Actual Maintenance Disbursements Using Month Dummy Model



Forecast vs. Actual Maintenance Disbursements Using Trig Function Model



Errors Produced by Current, Month Dummy, and Trigonometric Forecasting Models

Current	Average Monthly % Error	Average Monthly Forecast Error
July	-5.78%	-3,432
August	5.67%	1,859
September	15.13%	4,731
October	22.87%	9,189
November	-8.88%	-4,540
December	-7.44%	-3,680
January	25.89%	5,558
February	66.79%	5,271
March	19.61%	1,188
April	-0.92%	-909
May	-1.51%	-3,539
June	9.91%	1,388
Month Dummy		
July	-5.92%	-3,647
August	-0.61%	-788
September	11.04%	3,522
October	-3.14%	-1,522
November	1.19%	-733
December	-7.07%	-3,550
January	12.03%	2,780
February	35.38%	508
March	23.50%	2,474
April	11.28%	3,337
May	4.55%	-941
June	0.80%	-2,261
Trig Model		
July	-3.43%	-2,357
August	1.43%	208
September	6.73%	1,890
October	-7.12%	-3,488
November	3.27%	-28
December	10.54%	2,710
January	0.63%	-1,259
February	44.66%	2,492
March	11.87%	-303
April	-1.45%	-1,048
May	5.67%	-792
June	7.08%	134

APPENDIX E

Maintenance Expenditures Submodel
Forecast Template

Template for Use in Forecasting Maintenance Disbursements

The forecasting equation is embedded in column 2.

Month and Year	Forecast	Annual Budget/12	Trig 1*	Trig 2*	Non ag wage and Sal Empl*	Prime Rate Laged 18 Months	Election Year Dummy
Jul-94	65614.33	51050379.58	23075743.85	43446750.84	2958.63	6.00	1
Aug-94	63374.07	51050379.58	26422803.89	-1490646.70	2954.47	6.00	1
Sep-94	59152.18	51050379.58	22689894.96	-44937397.54	2985.81	6.00	1
Oct-94	55176.45	51050379.58	12877247.01	-43446750.84	3005.63	6.00	1
Nov-94	52794.04	51050379.58	-385848.89	1490646.70	3017.88	6.00	1
Dec-94	50193.84	51050379.58	-13545556.88	44937397.54	3026.52	6.00	1
Jan-95	38399.1	51050379.58	-23075743.85	43446750.84	2945.62	6.00	0
Feb-95	33158.99	51050379.58	-26422803.89	-1490646.70	2949.93	6.00	0
Mar-95	32928.45	51050379.58	-22689894.96	-44937397.54	2987.88	6.00	0
Apr-95	39266.97	51050379.58	-12877247.01	-43446750.84	3020.66	6.00	0
May-95	50395.29	51050379.58	385848.89	1490646.70	3039.73	6.00	0
Jun-95	61832.81	51050379.58	13545556.88	44937397.54	3070.86	6.00	0

* These are series that are generated previous to the forecast. The asterisk indicates hidden columns containing the data used to generate Trig 1, Trig 2, and Nonagemp.