Division of Research
Graduate School of Business Administration
The University of Michigan

May 1986

THE ACCURACY OF DATABASES FOR FORECASTING: A REVIEW AND EMPIRICAL STUDY

Working Paper No. 462

Gillian Rice
University of Michigan-Flint

Essam Mahmoud
University of Michigan-Flint

The Accuracy of Databases for Forecasting: A Review and Empirical Study

Abstract

The performance of any forecasting technique will always be related to the quality of the database to which it is applied. The increase in databases available and concerns about the accuracy of some of these databases and data sources means that evaluation of databases is a significant part of a forecasting system. This study addresses the accuracy of databases and reports the results of a survey of suppliers of databases for forecasting applications. Among the issues covered by the survey are database availability, types and formats of data, data sources used, and the approaches followed in maintaining data accuracy. In general, data accuracy appears to be sufficient but there are some practices followed by database suppliers which could produce unreliable information.

Keywords: Database, Accuracy, Forecasting.

INTRODUCTION

Advances in computer technology have meant a great increase in the capacity to communicate electronically and this has been accompanied by a growth in the availability of electronic data sources for use by businesses. By the early 1980's there was a growing maturity and acceptance of database systems. Most organizations were directing their major systems development toward on-line database systems (U.S. Department of Commerce, 1982). These developments have contributed to improved efficiency in market planning and forecasting.

It is important to remember, however, that the performance of any forecasting technique will always be related to the quality of the database to which it is applied. Hence, evaluation of databases is a significant part of the development of a formal forecasting system. Indeed, Lawrence (1983) stressed that one of the objectives for a quantitative forecasting system was assistance to the forecaster in "processing" past historical data.

This study addresses the accuracy of databases that are used in planning and forecasting by firms. There are questions of accuracy and reliability, for example, in considering how database vendors or suppliers check their data, and any forecasted variables that may also be provided. This paper reports the results of a survey of suppliers of databases for forecasting applications. Among the issues covered by the survey are: the data collection methods of the suppliers; the data sources used; the types of data collected; the formats in which data are available; and the approaches followed in maintaining data accuracy.

The paper is organized as follows. First is an overview of the use of databases for market planning and forecasting, with emphasis on the accuracy of databases. Next is a report of the survey of database suppliers with a detailed discussion of the survey's results. The conclusion considers the

implications of the results for forecasters and suggests future directions for research.

THE USE OF DATABASES IN MARKET PLANNING AND FORECASTING

There are two main uses of data in quantitative forecasting. First, data are used to determine the pattern of behavior of some variable based on historical observation. Data used in this way are incorporated into a time series forecasting methodology. Second, data can be used to provide the future values of independent variables included in a causal model. This usage often requires that these independent variables be forecast before they can be part of the input to such a causal model (Makridakis et al., 1983).

The kinds of information used for forecasting for marketing planning usually include activities and events both internal and external to the firm. Therefore, a wide range of variables must be forecast (Wheelwright and Makridakis, 1985). The first data source needed for sales forecasting is the firm's own record keeping system. Company records of shipments, financial records of manufacturing, distribution, and selling costs, and other data (such as price discounts, advertising expenditures, etc.) form the core of the management information database (Davidson and Ayers, 1982). These data are typically used in time series forecasting. External data sources are used more often in causal forecasting models. The major sources are the U.S. government (the largest source of demographic and economic information), periodicals and journals, trade associations, syndicated data sources and forecasting service organizations. In addition, Rice and Mahmoud (1985) discuss and list data sources that are particularly relevant for international business forecasting. There has been a tremendous increase in the number of published data sources available to forecasters, including many online

computer information services which have developed within the last decade.

The reliability and consistency of data may vary widely from source to source necessitating careful scrutiny of data and the procedures used to obtain data. There are many ways that data may be misleading if they are not evaluated carefully. Stewart (1984) stresses that evaluation should include issues such as the purpose of data collection; the potential biases on the part of the organization collecting the data; definitions of data units and concepts; the method of data collection; and the degree of consistency between the data and other sources. Specific criteria that are important in selecting a database from a commercial organization are coverage of the database, availability, accuracy, timeliness and the possible forecasting applications. If the database is to be accessed online, other considerations might be the ease of use by the forecaster, the type of data that can be accessed, the type of access, the degree of access allowed, the size of the database, the possibility of comparison with hard copy references, search options, search ranges and fields, spread features, print, display and storage options, documentation availability and customer services provided by the database supplier.

The emphasis of this study is on the criterion of accuracy of the data provided by sources external to the firm, in particular, the commercial organizations which provide and sell access to databases.

Data Accuracy and Forecasting

Data are evaluated for forecasting purposes for two reasons. One is that the forecaster must determine any patterns in the data in order to be able to select the most appropriate model for a data set. The second reason is that the accuracy of the forecast may depend on the quality of the data (Lackman, 1981). There is evidence that the accuracy of quantitative forecasting models

is not much affected by data errors and data revisions (for example, Armstrong, 1970; 1985). Fildes and Wood (1978), however, point out that often "it is not how you forecast but what you forecast" that is important. Fildes and Howell (1979) stress that although revisions to a "true" data series are unimportant, the choice of data can be quite critical. From a survey of Australian firms, Lawrence (1983) found that among the reasons for firms abandoning quantitative forecasting were the following: (1) the forecast user had jeopardized the success of the model by failure to maintain its database, and (2) forecast users complained of the large effort required to manage the forecast database to remove extraordinary events. While certain forecasting models may be able to deal differently with data problems such as outliers, the general problem of data quality--accuracy--affects all forecasting models equally. Hence, the issue is what level of accuracy is required for a particular forecasting situation. This will vary according to the importance of the management decision and the role of the forecast in that decision (Wheelwright and Makridakis, 1985).

Data Accuracy: The Evidence

One difficulty in assessing accuracy might be that not all databases are well documented and it may not be clear how data were collected or coded (Stewart, 1984). Generally, there is little information about the level of accuracy of published data sources. Wheelwright and Makridakis (1985) discuss an extensive study undertaken some years ago by Morgenstern (1963). This study examined the accuracy of governmental data sources and found that the National Income and Product Accounts (NIPA) consumption expenditures had a probable error range of + 10 percent to + 15 percent. If this error range is still applicable, with consumption expenditures at their current levels, the variable could be over- or under-stated by over $100 billion. In NIPA,

consumption is reported to the nearest $0.1 billion. Wheelwright and Makridakis comment that such reporting obviously can be misleading to the manager who does not examine the accuracy of the data source.

A Business Week article (May 6, 1985) discusses how many government statistics are becoming more volatile and unreliable. One factor contributing to this is declining budgets (Marketing News, February 28, 1986). Another factor is that the government's statistics are based on assumptions made years ago and which fail to capture adequately revolutionary shifts in the economy. Examples of these are the surging employment and output in the service sector, high-tech manufacturing growth, and the way imports are seizing a growing share of consumer and capital spending. While there is concern over the quality of government economic data, Parker (1982) argues that there is no easy way to measure this quality statistically. The NIPA estimates, for example, are constructed from hundreds of different sources, and the integration of these sources into the methodology requires a high degree of judgement on the part of government staff.

Although the private sector is a large provider of economic data, most information companies take raw data from the government, then reassemble, interpret and sell information. Without the government's ability to gather quality data, the ability of private companies to provide good economic information is constrained. Trade associations and market research firms gather data at a narrower industry level; these data have been found to be unreliable, however (Business Week, May 6, 1985).

Cateora (1983) commented on the problem of reliability in international data sources. Data from the U.S. and the highly industrialized nations are likely to have a higher level of accuracy than those from developing countries, largely due to the methods of collecting data. The Organization

for Economic Cooperation and Development (OECD) provides some of the most accurate data, but it has sometimes been criticized for adjusting forecasts to the official line of its member governments. In the case of the International Monetary Fund's "International Financial Statistics," however, while the data are provided by national sources, they are screened and reviewed before being included in the statistics. If the compilers have any reason to believe that the data provided by a national government are "doctored" they simply exclude them from their publication.

Despite the lack of evidence on the accuracy of data obtained from databases external to the firm, it should be noted that it is almost always less expensive to use these data than to conduct primary research. It is often possible to use data for ways in which they were originally not intended. Furthermore, when users are charged for information, there is an incentive to ensure quality and consistency on the part of the information providers. When using government statistics such as Gross National Product (GNP) as an input to a causal model, however, it might be preferable to base forecasts on a range of GNP assumptions, and likewise for other environmental variables (Jenkins, 1982). Fildes and Howell (1979) further suggest that if a forecasting model is unduly sensitive to the data on which it is based it must be used carefully since the magnitude and consistency of data errors are often unknown.

Investigating Data Accuracy

Because of the importance of data accuracy for forecasting and the limited knowledge about the accuracy of databases external to the firm, it is useful to investigate the accuracy of these databases. Given the obvious difficulties involved, the selected approach was a survey of database suppliers which included issues of data accuracy maintenance. Database

integrity is related to the database administration's responsibility for the accuracy of data and can be achieved through the use of validation checks, loggings, dumps, backup and recovery procedures, and auditing procedures (U.S. Department of Commerce, 1978). A major problem with data errors is that data enters the computer already contaminated. Some causes of bad data include; human error in initial recording or formatting of the data, faulty transcriptions of the data, falsified or tampered data, accidently omitted data and so on (McKeon and Buckanic, 1978). Optical disk technology, however, which allows documents to be entered in a visual form, requires fewer data entry operations and will reduce errors (Mortenson, 1984).

In this study's investigation of how database suppliers maintain the accuracy of their data, the following issues were examined: whether the database includes calculated or raw data; the sources from which data are collected; the method of data input; the frequency of data input; the method of checking accuracy; the frequency of accuracy checks; the primary source of errors; the method of correction once an error is detected; and the speed of correcting data errors. The following section discusses the methodology of the database supplier survey and the survey's results.

THE SURVEY OF DATABASE SUPPLIERS

Methodology

A comparison of the 1984 Database Directory and the 1984 Omni Online Database Directory revealed a list of 185 suppliers of historic and interactive databases for business forecasting use. The suppliers were located in the U.S. and ten other countries. A questionnaire was mailed to 137 U.S. organizations and 48 overseas organizations. The questionnaire sought information on the number of different databases supplied, the types of

business data included in the databases, the sources of information used and the frequency of use, the relative importance of a group of criteria for selection of different data sources (such as governmental, corporate etc.), the types of databases offered, data format and variables, the types of customer according to the area of business, and several details on the maintenance of data accuracy. The rate of response was 31 percent.

Discussion of Results

Types of Databases Offered by Responding Suppliers. The majority (94 percent) of the database suppliers offer source databases which contain numeric, textual numeric and full text information. These databases are therefore suitable for quantitative forecasting. Twenty-eight percent of the suppliers have reference databases which may also be used for qualitative and technological forecasting. For example, in technological forecasting the data used are specialized, often primary scientific reports, intelligence data and general information from publications. The data themselves are not of uniform importance in such situations. Rather, it is the way the data are processed to form future alternatives that is critical (Makridakis et al., 1983).

Most of the organizations supply more than two databases. The number of variables included in the databases ranges, for example, from eight for FISCAL (National Computer Network of Chicago) to 25,000 for CANSIM (online database of the Conference Board of Canada). Most organizations offer data on hundreds of variables.

The wide variety of databases available stresses the importance of selecting the database that has the right amount and type of information for forecasting purposes. The types of business data collected by the organizations in the study are shown in Table 1. Although some of the databases may provide information in the same subject area the number of

Insert Table 1 here

variables and the specific data provided may differ. The user, therefore, must examine a database and match his or her needs with the database offerings before subscribing to a particular database.

An important issue from the point of view of the forecaster is the format in which data are provided. It is crucial for forecasting that data be available in disaggregated form. Aggregation causes loss of information (Makridakis et al., 1983). In addition the levels of aggregation may be inappropriate and classifications may not match the forecaster's particular needs (Stewart, 1984). In the classification of demographic variables, for example, there are often no accepted definitions for the concepts measured. Wheelwright and Makridakis (1985) discuss data aggregation in financial statistics collected by various federal agencies. As each agency collects data relevant to a somewhat different population, it is generally impossible for a corporation using the statistics to sort out the differences in a meaningful way.

Therefore, in the context of database evaluation, it is more efficient to access data available at the most detailed level possible and then to aggregate them rather than to purchase data at an aggregated level and later discover they should be broken down. Of further importance is the fact that the process of aggregating data was classified as one of the common sources of data error by Morgenstern (1963).

It is encouraging that almost three-quarters of the respondents supply raw numerical data, as shown in Table 2. Another advantage of access to disaggregated data is that they allow the forecaster to explore more easily

Insert Table 2 here

the characteristics of the data, in particular, the pattern of the data. Graphics, provided by 30 percent of the database suppliers, may also be useful here. Because different forecasting methods vary in their ability to identify different types of patterns, it is important to match the presumed pattern(s) in the data with the appropriate technique (Wheelwright and Makridakis, 1985). This may have implications for the accuracy of a particular technique.

Sometimes it is also valuable for forecasters to access well-organized statistics or forecasts and these are provided by 58 percent and 10 percent of the organizations, respectively.

Data Sources. When attempting to assess accuracy, it is useful to examine the data sources used by suppliers to compile their databases. The number of organizations from which suppliers collect data varies somewhat. Over 50 percent of the suppliers collect data from less than 100 organizations, and slightly less than half from over 100 organizations. The frequency of use of different kinds of organizations is illustrated in Table 3. U.S. government data and data from other public sources and individual corporations are used most consistently.

In order to understand how database suppliers evaluate the data sources they use, they were asked to rate the importance of a number of selection

Insert Table 3 here

criteria that might be relevant to the choosing of the data sources. The results are contained in Table 4. The criterion of data availability is rated the most important in all cases. Only if data are available can they be used

to compile a database. Reliability is also seen as an important criterion for the selection of data sources by database suppliers. This is encouraging for forecasters as the suppliers are implying that they are concerned about data reliability. The criterion of reputation, which may be related to perceptions of accuracy, also appears to be of some importance for most of the suppliers. Completeness of data is a criterion of more importance in the case of federal

---

Insert Table 4 here

---

government or corporate sources that it is for the other sources. Indeed federal government data are used often because of their completeness. On the other hand, incompleteness of international data is often a problem for users of these data.

The least important selection criteria in general are the organization of data and high quality documentation. The database supplier, as a compiler of raw data, probably provides its own documentation and, in compiling the data, organizes them according to users' needs. The actual processes of compiling and organizing data may have some effect on the accuracy of the database.

Maintaining Database Accuracy. In a survey of this type, the evaluation of database accuracy was inevitably restricted to measuring accuracy in some indirect fashion. Also, rather than investigating suppliers' perceptions of the overall accuracy of their databases, the emphasis was placed on learning how suppliers maintained the accuracy of their databases. A number of variables thought to be related directly to database accuracy were identified and examined. These were the method of data input, the method of checking accuracy, the frequency of accuracy checks, sources of error, speed and method of correction, and the method of dealing with outliers.

With respect to the data input method, 28 percent of the database suppliers input data instantaneously, 47 percent use a batch input method, and 64 percent input data by a batch update method. Notice that several organizations use more than one input method. The particular input method used has implications for the timeliness of data for forecasting. The lack of timely data can affect the appropriateness of a forecast for a certain time period which could have ramifications for the accuracy and success of market planning. It is therefore of some concern that only 28 percent of the suppliers input data instantaneously. In the case of using databases provided by suppliers implementing batch or batch update methods, the forecaster would need to know when the batch processing and updating took place. Forecast model accuracy could then be checked as the actual data become available. Data collected and published on a timely basis are of greatest value to the forecaster (Levenbach and Cleary, 1981).

Future research could investigate whether the frequency of errors differs according to the method of data input. This study found that transcribing was the most common source of error. As mentioned earlier, however, the number of transcribing errors should fall as the use of optical disk technology becomes more widespread. Table 5 shows that "unreliable source" was the second most common source of error. This result reinforces the importance of learning about the data sources that the database suppliers utilize.

Insert Table 5 here

Almost 20 percent of the database suppliers responding to the survey indicated that they do not check the accuracy of the data they receive (see Table 6). Most suppliers also do not apply sophisticated methods such as

using a special editor or analyst or by mathematical or statistical testing. If an error is found, only 48 percent of the database suppliers reported that an expanded sample would be taken to search for additional errors.

Insert Table 6 here

Furthermore, over 20 percent of the suppliers do not recheck the accuracy of their database information. While Table 7 shows that 44 percent of the suppliers recheck accuracy continuously, 14 percent do so only on a monthly basis or at even longer time intervals. This could be problematic for the forecaster trying to build a model using bi-weekly or daily data. With respect to error correction, about half of the suppliers correct errors

---

Insert Table 7 here

---

immediately and most do so within 48 hours. Table 8 illustrates that 48 percent make corrections to errors with other updates through the batch mode. There is, therefore, a certain probability that a forecaster is accessing uncorrected data. It should be noted that database suppliers may check and/or correct data at different times for different variables. It would be valuable for a forecaster to have access to such information pertaining to a database

---

Insert Table 8 here

---

being used. The results of this study suggest that there is an opportunity for many database suppliers to improve the accuracy of their databases.

How database suppliers deal with outliers found during checking procedures can be an important issue for a forecaster. Outliers most often

create the greatest difficulty for the forecaster since they are generally unexplainable. A thorough understanding of the source and accuracy of the data is required to recognize the true importance of irregular data (Levenbach and Cleary, 1981). The results pertaining to dealing with outliers are shown in Table 9. From a forecasting perspective, data should be processed into the database "as is", (once verified), to enable the forecaster to make his or her own decisions about data adjustments. For example, Levenbach and Cleary (1981) stress that if someone is attempting to build a model to explain past results, the past data probably should not be adjusted. They may be useful in understanding how extreme or unusual effects affect the process of generating data. If the forecasting task is extrapolation, on the other hand, it is important to negate outliers by replacing them with more "typical" values. An extremely disturbing finding, therefore, is that 21 percent of the database

---

Insert Table 9 here

---

suppliers report that outliers are deleted from the data. Only one supplier stated specifically that his organization confirmed the outlying data with the source for possible immediate correction. In this way the supplier can determine whether the outlier is caused by an input error or whether it represents an extraordinary event. The question can be raised of whether the responding suppliers are fully aware of the impact of outliers for forecasting and modelling purposes.

CONCLUSION

The purpose of this research study was to investigate the accuracy of databases used in forecasting and marketing planning. In the context of the

study, accuracy of data is related to the type, format, and timeliness of data and how database accuracy is maintained. The large number of databases and included variables suggests that database selection is a significant decision for a forecaster.

From a forecasting viewpoint, the data format of the respondents is good with a large proportion of them supplying raw numerical data. Calculated statistical data are also widely available. The timeliness of data may be of some concern to forecasters depending on the type of model and time lags used. Foreign time lags are often harder to control, but domestic time lags should be as short as possible in order to produce the timeliest information for forecasters and planners. The overall findings of the survey with respect to maintaining database accuracy suggest that although general data accuracy is sufficient in most respects, there are some practices which could produce unreliable information. It is important that users of databases are aware of this. Forecasters must routinely check data they receive from various databases. For example, this can be done by comparing a random sample of data from one source with corresponding data from another source or the original source. As discussed above, it is useful for forecasters to know how and from what sources a database supplier collects data, the methods and frequency of accuracy checking and how data errors are corrected. Also, all database suppliers should check the accuracy of their data and should remain sensitive to the reliability of their data sources. In particular, the practices of dealing with outliers should be reviewed and, where possible, error correction procedures should be immediate and not batch.

Further studies of database accuracy could take more specific approaches, by investigating and objectively testing the accuracy of particular variables such as price indices and demographic data with regard to different database

suppliers. Database accuracy is an important area of study because successful marketing planning and other business decisions depend on accurate forecasts which, in turn, depend on accurate data. The results of this, and future research, will provide information valuable to the forecaster in the database selection decision.

Table 1.

Types of Business Data Collected by the Organizations in the Study

		Respondents (n = 52)	
		Number	Percent
1.	Price Indices	27	51.9
2.	Stock Prices	25	48.1
3.	Asset Holdings	24	46.2
4.	Historical Sales Figures	21	40.4
5.	Profit Figures	20	38.5
6.	Demographic Data	20	38.5
7.	Earnings Per Share	19	36.5
8.	Interest Rates	19	36.5
9.	Liabilities	19	36.5
10.	Price Levels	18	34.6
11.	GNP	17	32.7
12.	Employment/Unemployment	17	32.7
13.	Imports	16	30.8
14.	Exports	15	28.8
15.	Composition of Labor Force	14	26.9
16.	Projected Sales Figures	13	25.0
17.	Recent Mergers	13	25.0
18.	Payroll Figures	10	19.2
19.	Tax Data	9	17.3
20.	Government Contract Specifications	8	15.4
21.	Commodities	6	11.5
22.	Exchange Rates	2	3.8
23.	Current Events	2	3.8
24.	Patents, Trademarks	2	3.8
25.	Management Literature	2	3.8
26.	Banking Data	2	3.8
27.	Future Prices	2	3.8
28.	All Business Information Published in Daily Press of Historical Interest	4	7.7

Table 2.

Format in Which Data are Provided

		Respondents (n = 50)	
		Number	Percent
1.	Raw Numerical Data	37	74.0
2.	Calculated Statistical Data	29	58.0
3.	Text	26	52.0
4.	Models	20	40.0
5.	Reports	19	38.0
6.	Graphics	15	30.0
	OTHERS		
	Forecast	5	10.0
	Abstract	2	4.0
	Bibliographic	2	4.0

Table 3.

Sources of Information Used by Database Suppliers

According to Frequency of Use (n = 50)

Source	Consistent Use		Occasional Use		Rare Use	
	Number	Percent	Number	Percent	Number	Percent
U.S. Gov't Organization	28	56	6	12	3	6
State/Local Government	18	36	5	10	4	8
Other Public Sources	29	58	3	6	2	4
Individual Corporations	25	50	4	8	1	2
International Sources	18	36	8	16	--	--
Journals	22	44	3	6	1	2

Table 4.

Data Sources Used by Database Suppliers: Mean Ratings According to Selection Criteria (n = 38)

Selection Criteria	Fed. Gov't		State/Local Gov't		Other Public Sources		Indiv. Corp.		Internat. Sources		Journals	
	Mean Score*	Rank	Mean Score	Rank	Mean Score	Rank	Mean Score	Rank	Mean Score	Rank	Mean Score	Rank
Reliability	2.25	(2)	2.75	(2)	3.00	(2)	3.00	(3)	3.00	(2)	2.67	(3)
Reputation	3.83	(5)	3.50	(3)	3.80	(4)	3.88	(5)	3.40	(3)	3.00	(4)
High Quality Documentation	4.88	(7)	4.50	(6)	4.45	(6)	4.28	(6)	3.83	(5)	3.20	(5)
Minimal Lag Time Between Data Generation & Data Reception	2.83	(4)	3.70	(4)	3.25	(3)	2.50	(2)	3.40	(3)	2.50	(2)
Organization of Data	4.12	(6)	5.00	(7)	4.70	(7)	4.44	(7)	5.25	(7)	3.78	(6)
Data Availability	2.00	(1)	2.00	(1)	1.95	(1)	1.67	(1)	2.66	(1)	2.25	(1)
Completeness of Data	2.75	(3)	4.25	(5)	4.00	(5)	3.17	(4)	4.38	(6)	4.11	(7)

* Each score represents the mean rating score for a particular data source on a particular criterion. "1" represents "most important criterion" when selecting a particular data source. "7" represents "least-important criterion.

Table 5.

Sources of Error*

Source of Error	Responses (n = 47)	
	Number	Percent
Transcribing Error	26	55.3
Mathematical Error	2	4.3
Data Omission	7	14.9
Unreliable Source	12	25.5

* Respondents were asked which source of error accounted for the largest portion of errors.

Table 6.

Methods Used to Check Accuracy

Methods	Respondents (n = 43)	
	Number	Percent
Does Not Check Accuracy	8	18.6
Random Sample Review	14	32.5
Cross-Check With Primary Data Source	18	41.9
Comprehensive Review	13	30.2
Editor/Analyst	2	4.7
Testing Using Mathematical Relations and Standard Deviation	1	2.3

Table 7.

Frequency of Accuracy Rechecks

Frequency	Respondents (n = 50)	
	Number	Percent
Does Not Recheck Accuracy	12	24.0
Continuously	22	44.0
Daily	7	14.0
Weekly	2	4.0
Monthly	2	4.0
Quarterly	4	8.0
Annually	1	2.0

Table 8.

Error Correction Procedures Used

Procedure	Respondents (n = 52)	
	Number	Percent
Correction Made With Other Updates Through Batch Mode	25	48.0
Immediate Update	30	57.7
Contact Users Through Computer Bulletin	9	17.3
Return Poor Data	1	1.9
Once Online, Corrections Not Made Unless Errors Very Severe	2	3.8

* Note that several organizations use more than one error correction procedure.

Table 9.

Method of Dealing With Outliers

Method	Respondents (n = 47)	
	Number	Percent
Process Data "As Is"	18	38.3
Delete From Data	10	21.3
Statistically Adjust Data	15	31.9
Confirm Data With Source for Immediate Correction	1	2.1
Various Complex Methods	3	6.4

REFERENCES

Armstrong, J.S., 1970, An application of econometric models to International Marketing, Journal of Marketing Research, 7, 190-198.

Armstrong, J.S., 1985, Long Range Forecasting: From Crystal Ball to Computer, Second Edition (John Wiley and Sons, New York).

Business Week, 1985, Does anybody really know how the economy is doing? May 6, 128-132.

Cateora, Philip, R., 1983, International Marketing (Richard D. Irwin, Inc., Homewood, Illinois).

Davidson, Timothy A. and Jeanne L. Ayers, 1982, Selecting and Using External Data Sources and Forecasting Services to Support a Forecasting Strategy, in: The Handbook of Forecasting: A Manager's Guide, Spyros Makridakis and Steven C. Wheelwright eds., (John Wiley and Sons, Inc., New York) 469-483.

Fildes, Robert and Syd Howell, 1979, On Selecting a Forecasting Model, TIMS Studies in the Management Sciences, Vol. 12 (North-Holland, Amsterdam) 279-312.

Fildes, R. and D. Wood, 1978, Data and the validity of forecasting models, Journal of Enterprise Management 1, 97-106.

Jenkins, Gwilym M., 1982, Forecasting in Organizations, Journal of Forecasting 1, 3-21.

Lackman, C.L., 1981, Forecasting information systems: The data base module, Journal of Systems Management, January, 29-35.

Lawrence, M.J., 1983, An Exploration of Some Practical Issues in the Use of Quantitative Forecasting Models, Journal of Forecasting 2, 169-179.

Levenbach, Hans and James P. Cleary, 1981, The Beginning Forecaster (Wadsworth, Inc., Belmont, CA).

Makridakis, Spyros, Steven C. Wheelwright and Victor E. McGee, Forecasting: Methods and Applications Second Edition (John Wiley and Sons, New York).

Marketing News, 1986, Budget cuts could stem flow of government data, February 28, 3.

McKeon, Joseph M. and William Buckanic, 1986, Legal Risks, Journal of Systems Management, January, 28-31.

Morgenstern, O., 1963, On the Accuracy of Economic Observation, (Princeton University Press, Princeton, N.J.).

Mortenson, Patricia K., 1984, Dealing with Data, Best's Review (Life/Health Insurance Edition), 85.

Omni Online Database Directory 1984, 1985, (Collier Books, New York).

Parker, Robert P., 1982, The Quality of the U.S. National Income and Product Accounts, Review of Public Data Use, 10, 1-8.

Rice, G. and E. Mahmoud, 1985, Forecasting and the Data Base: An Analysis of Databases for International Business, Journal of Forecasting, 4, 89-97.

Stewart, David W., 1984, Secondary Research Information Sources and Methods, Applied Social Research Methods Series Volume 4, Sage Publications Inc..

The Database Directory, 1984, (Knowledge Industry Publications, New York).

U.S. Department of Commerce National Bureau of Standards Special Publication 500-28, 1978, Database Administration: Concepts, Tools, Experiences, and Problems, March.

U.S. Department of Commerce National Bureau of Standards Special Publication 500-92, 1982, Data Base Directions Information Resource Management-Strategies and Tools, September.

Wheelwright, Steven C. and Spyros, Makridakis, 1985, Forecasting Methods for Management Fourth Edition (John Wiley and Sons, Inc., New York).