Simulate Your Way To Exceptional Forecasts

Robert M. Duncan, CFPIM, Duncan Consulting, and
John A. Estep, CFPIM, E/Step Software Inc.

This is a case study of using a new forecast simulation facility to achieve higher forecast accuracy, improved customer service, and reduced inventory. It includes a description and demonstration of the simulation facilities and shows how each benefits the forecasting process. The improvements resulting from using this system are presented.

Used properly, forecasting software handles the routine situations and identifies the exception items for closer scrutiny by the forecast analyst. But it is not enough simply to identify the SKUs requiring attention. There are many things today's powerful desktop computers with good color graphics can do to present the exception items in a way which both facilitates the analyst's detailed study while also saving time.

This presentation describes and demonstrates (using a computer and color projector) a variety of interactive forecasting tools and techniques—including both classical theory and heuristics—showing the situations where each is applicable. The demonstration uses data from actual operations, and the presentation concludes with a discussion of the results of using the facilities over the past year.

1. Introduction

By studying the source of a word, often one can learn its proper use. It has been said the etymology of "forecast" comes from sport: "fore" is a golfing term meaning "watch out" and "cast" comes from fishing and means "throw out". After several years of running an unsuccessful forecasting implementation and after a year and a half of a successful one, I have learned that forecasting is not a passive black box with history in and forecasts out. Rather, the success comes from "watching out" (exception reporting) and "throwing out" (simulation—as in casting out a new bait to see what is pulled in).

2. GE Silicones

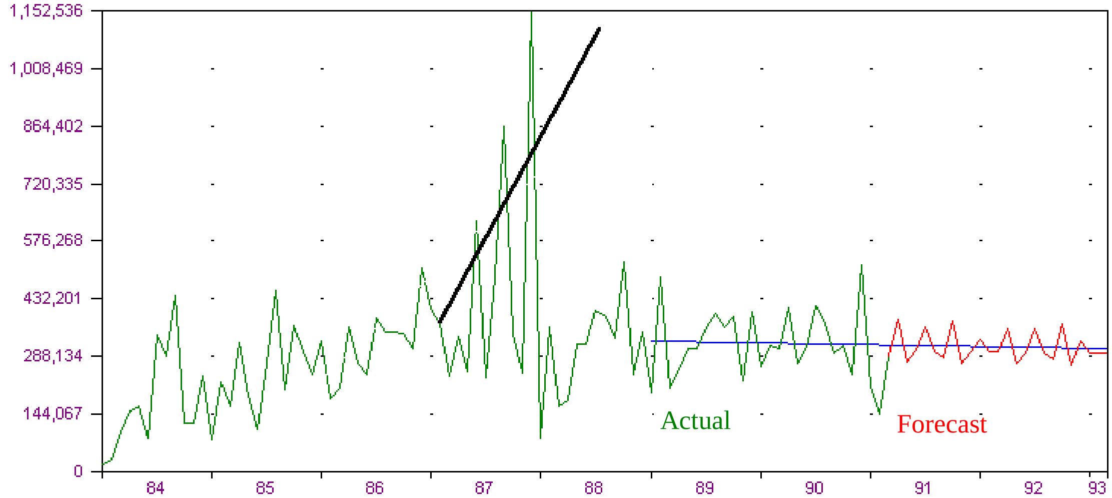
GE is one of the world's largest suppliers of silicone products. The company holds over 1,500 patents and manufactures over 4,000 silicone-based products. Silicones are silicon-based organic chemicals that exhibit the usefulness of organic chemicals and the strength of inorganic chemicals. Silicones exhibit paradoxically opposite properties. They can be used to add elasticity or rigidity; make things stick or slip; create foam or eliminate foam; serve as electrical conductor or insulator; and they can work at low or high temperatures.

The products are fluids, resins, and elastomers that are sold to manufacturers and consumers. Other products such as emulsions, greases, pressure sensitive adhesives and dispersions are derived from these basic groups. The applications are virtually endless, but center around the building and construction, industrial, automotive, electrical/electronics, aviation/aerospace, cosmetics, appliance and consumer hardware markets. GE Silicones is a highly vertically integrated business that turns "sand into rubber".

3. Lessons Learned From Failure

Forecasting for GE Silicones was originally a monthly calculation by a home-grown mainframe black box system. Identifying and reviewing exception items was difficult and seldom done. Furthermore, while the system looked for trends and seasonality, it used the outdated base index method. Also, while the company uses a 5-4-4 calendar for reporting purposes, the system performed no calendar normalization on the demand data. We did compute safety stocks based on forecast error, but using the obsolete mean absolute deviation (MAD), rather than the more accurate standard deviation. All of these things added up to inaccurate and insufficiently monitored forecasts.

A dramatic example of this is the SKU (stock keeping unit: a part at a location) shown in Figure 1. At the end of 1987 the system decided on a model which included a huge trend, partly as a result of successful promotions which generated record demand for December. Based on the trend, the forecast for January was even higher. Meanwhile actual demand for January was very low (because everyone stocked up at the December price). This increased the forecast error which in turn caused the system to recommend a higher safety stock. In short, while actual demand was low, both the forecast and the safety stock were high and increasing. The resulting high inventory was not just undesirable from a financial point of view, it meant that expensive capacity was tied up producing the wrong product. Furthermore, service for other products—which were needed—suffered.



Our efforts to improve the situation triggered some of the classical problems seen in production and inventory management. We permitted marketing to override the computer-generated forecasts. Marketing product managers, in an effort to insure they would get some of the production for their customers, over forecasted their own products. They never decreased the forecasts for over-forecasted products, however. And why should they? Manufacturing—not marketing—had responsibility for the finished goods inventory. The forecasts became very biased.

This resulted in heavily front-loaded master schedules. A "bow wave" of planned manufacturing orders was sitting on the other side of the planning fence. The production schedulers, constrained by capacity from scheduling to cover the forecasts, were asking "What is the real demand?" So they scheduled what they guessed was actually needed—and they became the real forecasters. Of course they got a lot of telephone calls from marketing when they guessed wrong—and marketing became the real expediters! The net result: service suffered and inventory climbed.

4. Organizing For Success

We achieved success by addressing three areas: organization, software, and focus. First, we created a Materials and Logistics department and gave it the responsibility for customer service and finished goods inventory. The new organization reports to the general manager at the same level as manufacturing, finance, marketing, and sales. The new department is divided into Customer Service, Purchasing, Inventory Control, Scheduling, and Planning and Analysis. Within the Inventory Control group, five finished goods inventory planners are responsible for maintaining the right inventory levels and customer service.

Second, we installed the "Finished Goods Series" (FGS) software from E/Step Software Inc. of Yakima, Washington. FGS is a PC-based demand forecasting and inventory planning package which, while it runs on the PC, handles large volumes of data and is designed to exchange information with mainframe-based order entry, scheduling, and other systems through a series of flexible interfaces. While FGS is set up to handle large numbers of SKUs routinely (i.e., hands off), it has excellent facilities for identifying and reviewing those exception items which require something other than routine handling. The package includes a recently-developed simulation module which allows us to view the results we would achieve with various strategies for handling exceptions—but before we actually make the changes. In this way we can compare a number of scenarios and then pick the best one. While there are literally dozens of ways to quantify forecast error (e.g., x% of the SKUs had less than y% error, etc.), FGS lets us measure forecast error in terms which are meaningful to the business: safety stock levels in units and/or dollars. If our customer service objectives dictate a safety stock level which is so high that the business is unprofitable, then we know our forecast errors are too large, and by how much.

Third, we used Pareto analysis to identify the few critical items which make up the majority of our business. For us, about 20% of the products account for 80% of the orders. We set FGS' control limits for these items to much tighter tolerances than we used for the rest of the items. This enables us to focus our attention on the few items where it is merited and where there are substantial benefits associated with any improvements. In contrast, the more typical approach to forecasting is one where the system produces a foot high stack of paper every month, and it is up to the planner to try to find the problems.

5. Initialization And Process Overview

We loaded about five years of demand history into FGS and let it find the optimum model for each SKU. We could have loaded less data, but many of our products exhibit seasonality and using only two or three years of history would have given us a weaker foundation upon which to build. If an item exhibits seasonality, then—in a sense—five years is only five observations, which is little enough. We were able to handle the problem of our 5-4-4 calendar by loading our calendar period-ending dates (past and future) into FGS and letting it perform the normalization automatically. This means that any seasonality it detects is true seasonality, and not just caused by our accounting calendar.

Each month we load the demand for the month just completed. Next we "filter" the demand to check for reasonableness. We compare the forecast and actuals with the marketing intelligence to see whether it helped or hurt. Then we let FGS use the computed error to revise the forecasts and models using adaptive smoothing. This process also identifies exception items which we review and, in some cases, modify. The FGS Simulation Facility is the key to examining the exceptions. We use it to try one or more alternative scenarios before deciding on the changes. Finally we run summary and detailed reports, including interfaces to the mainframe which transfer the new forecasts and safety stocks.

6. Why Not To Refit

One frequently-asked question is, "Why revise? Why not just fit new models every month?" The quick answer is that refitting takes more computer time and the results are not as good as with revising. But to consider the subject in a little more detail, our products—as do most companies' items—exhibit a fundamental stability most of the time. By that I mean that if a forecast model truly represents the underlying demand for a product, then that model will be effective for some period of time. It's not that one model will work one month, and a totally different one is required next month. There are most certainly changes, but they are usually changes in degree, not in kind. You would not see a product with a level, trend, and quarterly seasonality go to just a level and trend next month and to semiannual seasonality the month after that. So, the purpose of the model fitting process is to determine the most appropriate model, whether using classical statistics, clever heuristics, or even manual intervention.

The adaptive smoothing and error tracking that occurs in the forecast revision process has two purposes. The first is to make those changes in degree (but not in kind), which keep the model up-to-date with reality. This handles situations such as a trend which is gradually flattening or seasonality which is becoming less pronounced as we expand into new markets in what used to be the off season. The second purpose of the revision process is to identify those items where the chosen model is suspected of being no longer appropriate (i.e., changes in kind, not just degree).

This is vital to us: the system tells us that what used to work for an SKU does so no longer. With that knowledge in hand we can investigate the cause of the change, be it new competition, product changes, or whatever. This is counter to the belief, which some hold, that one should try every model on every SKU every month; but we know our approach works for us.

7. Filtering The Demand

The Demand Filter Report is our defense against order entry errors corrupting our forecasts and can also help spot trend changes.

Product	Fcst Demand	Actual Demand	Filter Sens.	-----Filter---- Min	Max	Dollar Error
C21	110,025	17,196	3.0	35,388	184,662	40,724
A15	12	770	4.5	-30	54	32,540
A42	96	27	3.0	33	159	16,555
E11	954	13,000	4.5	-8,356	10,264	12,327
M24	5	20	4.5	-9	19	9,684
F16	22	197	4.5	-68	112	2,006

Figure 2 — Demand Filter Report

As the sample in Figure 2 shows, we have given each SKU a filter sensitivity, expressed in standard deviations. The smaller the sensitivity value, the more likely the SKU is to show up on this exception report; i.e., the more scrutiny it receives by the planners. The sensitivity is translated into a minimum and maximum filter. Actual demand falling outside these values (either too high or too low) causes the SKU to appear on this report. The SKUs on the report are sorted in order from largest to smallest error in dollars. That way, if we do not have enough time to review them all, we know we have looked at the most important ones first.

8. Using Orders Booked

We have a Forecast Alert Report (on the mainframe, although it could be done in FGS instead) which identifies all products where orders booked beyond the current month already exceed the forecast plus safety stock. Even though we produce to forecast, if the customer is considerate enough to give us advance warning on a very large order for a make-to-stock product, we shouldn't miss it. In such cases we manually override the forecast (entering "marketing intelligence"). This gives the planners a warm feeling that they can plan to the forecast.

Given our past experience, however, I was worried about forecast bias. With an unbiased forecast, orders should exceed the forecast half the time and forecast should exceed orders the other half. But since we are only "biasing" those products that (1) exceed about two standard deviations (based on our service objectives), and (2) are ordered at least one month in advance, this happens very rarely—about two to five products per week. Besides improving customer service (covering the percentage of service that our safety stocks are not designed to cover), the more accurate forecasts contribute to a safety stock reduction! Monitoring the orders booked in this way can also forewarn of trend changes.

9. Evaluating Marketing Intelligence

In our FGS training we were taught that we should use marketing intelligence when we know something the system does not. If we are simply guessing, we should not use marketing intelligence because FGS can guess better than we can. This means that forecast overrides should always improve the forecast, but experience teaches that we all make mistakes occasionally. Learning from our mistakes (and from our successes) is crucial to improving the forecasting process. The Marketing Intelligence Evaluation Report is our vehicle for learning to generate more accurate marketing intelligence.

Product	Statistical Forecast	Adjusted Forecast	Actual Demand	Error Diff	Improvement $ Value
G54	14	10	9	+4	+17,464
C17	42	52	121	+10	+3,243
J63	10	13	13	+3	+2,229
B48	19	15	26	-4	-2,189
A67	16	27	20	-3	-4,004
D25	9	16	6	-7	-9,259
Net Dollar Improvement					+7,484

Figure 3 — Marketing Intelligence Evaluation Report

In Figure 3, the Error Difference column is the difference between what the error would have been using the statistical forecast and what it was with the adjusted forecast. The error differences are weighted by standard cost and the resulting dollar value of improvement is used to sort all the SKUs which had marketing intelligence last month. Here we made the forecast better for half the items (total improvement $22,936) and worse for the other half (total improvement $-15,452) for a net improvement of $7,484. For the first three SKUs we can give ourselves a pat on the back, and see if other situations would also benefit from using the logic we employed there. The last three SKUs show us where we have room for improvement, which is also valuable—though less pleasing—information.

10. Identifying Exceptions

The revision process identifies the SKUs for which there is some degree of suspicion that the model in use may no longer be appropriate: that is, it looks for bias in the forecasts. The method used to detect bias is the parabolically-masked cumulative sum of errors technique [Reference 1]—quite a mouthful, but it does a good job for us. The SKUs are flagged in the database via what amounts to an index. This means we can use the index to trigger any reports or processes we want, rather than just getting a printout and having to reenter the part numbers when we want to take action.

We use the FGS Report Generator to display a list of the exceptions, showing part number, planner, forecast level, date of first demand history, modelfit exceptions, next month forecast, forecast error, last month forecast, and safety stock in dollars (one line per SKU). We print the list sorted by planner, and within that by descending dollar value of the safety stock. This is my favorite report as it is our early warning of changing demand patterns and we use it to focus our attention and direct our actions.

In the beginning, I set the tracking signal sensitivity for all SKUs to four standard deviations (just like the text books recommend) and got two pages of exceptions. That was all we had time to deal with. A couple of months later as we learned to use the tools better and the forecasts improved, the list dropped down to one page; and then to a half page. We tightened the sensitivity to three standard deviations and the list grew to two pages again. Bear in mind that as we tighten the sensitivity, we are doing a better job of monitoring the forecasts, and the exceptions generated are based on either smaller errors or shorter duration or both. Another way of looking at it is that tighter limits mean there is a greater probability that an exception is spurious and requires no action, but it does mean we can catch changes earlier. After a while we tightened the sensitivity again to 2.5 standard deviations and eventually, with further improvement, to our current 2.2. The greater sensitivity makes it possible to identify trends shortly after they change, which further serves to improve our forecasts.

This process cannot happen all at once, it has to happen over time. It is the continuing improvement in our ability to forecast which allow us to tighten the sensitivity. Otherwise, we would just get a longer and longer list of exceptions. Also, not all products are set to this tight sensitivity—just the ones that merit the close scrutiny. The others are set to wider limits and consequently require less attention.

11. Using Simulation To Handle The Exceptions

Once a month I meet with the inventory planners for one to three hours each to review the exceptions and any questionable forecasts they have identified. We review the exceptions in descending order by safety stock value, which means we start with the products which have the greatest potential inventory savings. We look at the products using the FGS Simulation Facility, with the goal of generating forecasts that—based on our knowledge of the products and markets—look correct on the graph (level, trend, seasonal pattern, etc.). In addition, we monitor the safety stock in attempt to achieve reductions (via error reductions, not service reductions) wherever possible.

We continue this process until we run out of either exceptions or time. We don't always change every item reviewed, sometimes we wait until next month to see if the exceptional demand continues or returns to its prior pattern. When we do make changes there is usually a significant reduction in safety stock. Sometimes, however, the safety stock goes up (indicating a less accurate forecast based on demand history), and sometimes it goes up when the forecast has gone down. In these cases we know that, because the forecast will eventually be proved to be more accurate, over time the safety stock will come down. But in the meantime the increased safety stock protects our customer service until our assumptions are proven correct. This lets us boldly modify the forecast while still being conservative when it comes to customer service.

In the next sections we'll look at some of the typical exception situations and how we handled them.

12. Product Life Cycle

Demand patterns are most easily identified by viewing a graph of the history, which the simulation provides us. This first example in Figure 4 is a classic illustration of the life cycle of a product (grow, mature, decline, and hang on forever).

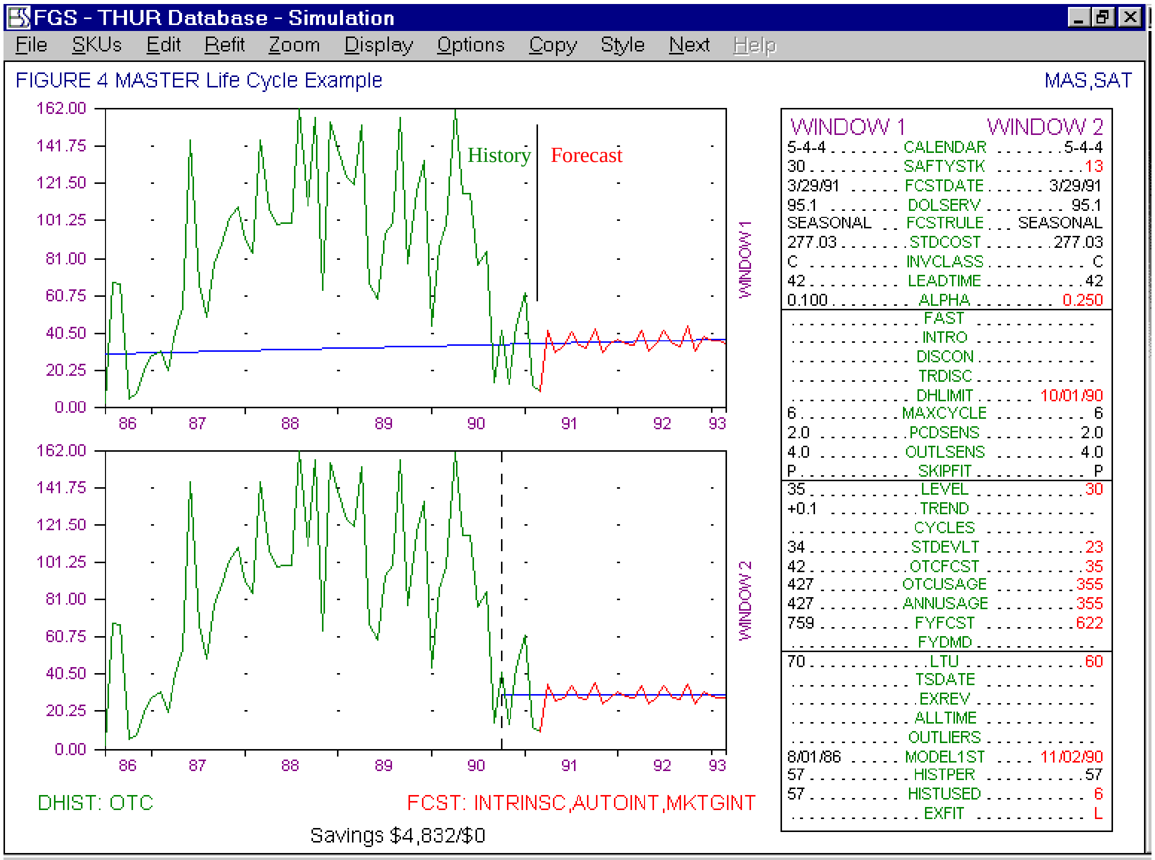


Figure 4 – Life Cycle

The eye can readily pick up this pattern in the graph, and the planner—knowing the story behind it—handles it by setting a demand history (DH) limit. This limits the history considered when refitting the model. In this example we limited the history to the last six months, which reduced the forecast for next month from 42 to 35. Given that the actual was 11, one could well argue that it should have been lower still, but we played it conservatively and there were further opportunities in the ensuing months to review it again. Also the system moved the item to fast smoothing which means it will automatically drift down more rapidly if the demand continues to fall. In the meantime our action reduced the safety stock by 44 percent or $4,832—a tidy savings.

13. Demand For A Product Takes Off

Figure 5 shows a product for which the tracking signal flagged a trend change that the adaptive smoothing was moving too slowly to catch. Adaptive smoothing reduces forecast nervousness (overreacting to noise), but it always lags a change as a result. The tracking mechanism caught the explosive growth in the last couple of months and merely refitting the model sufficed to correct the situation. As a result both the forecast level and trend increased. As a check: next months forecast increased from 205 to 258; the actual was 250. Notice also that the lower error in the new model reduced safety stock by 18 percent.

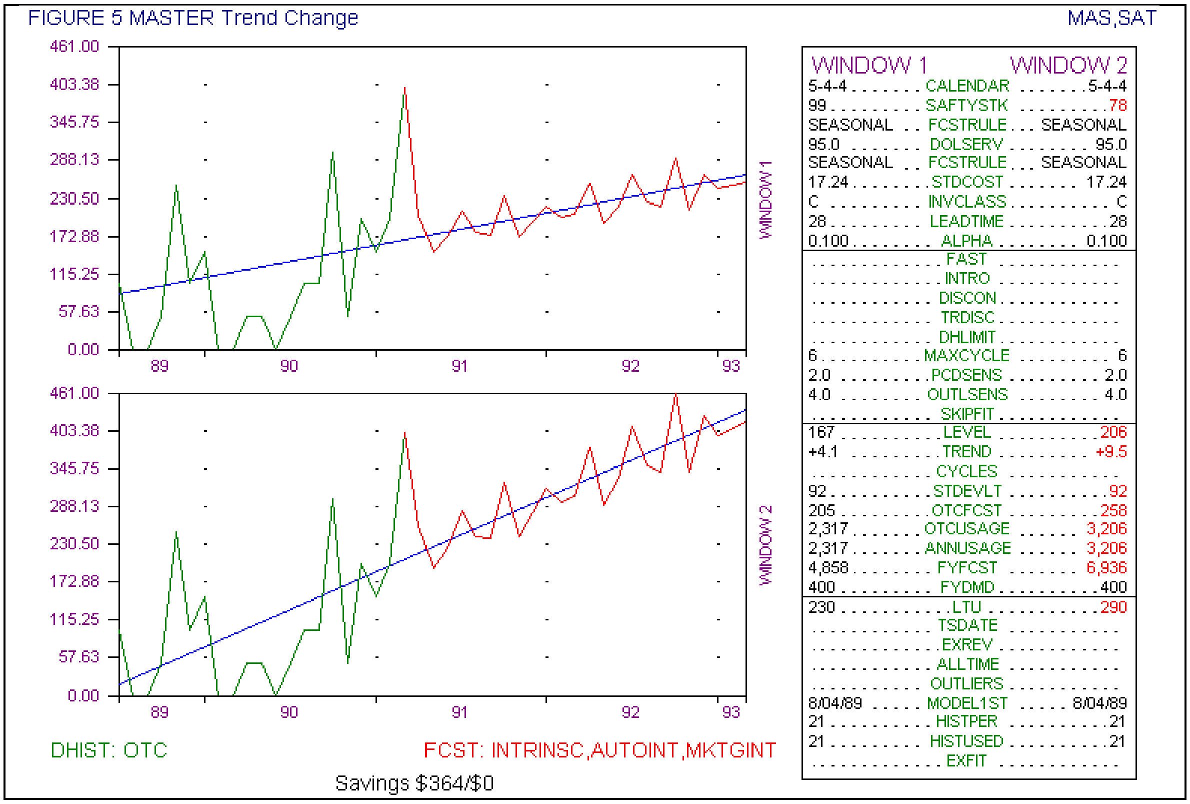


Figure 5 – Trend Change

14. Multiple Changes: DH Limit And Outliers

The product in Figures 6 and 7 benefited from two changes, based on the planners knowledge of the situation. This product was picked up on both the Demand Filter Report and the Tracking Signal Report. Since the beginning of 1988, the demand has not been as volatile as in prior years. We set a demand history limit for January 1, 1988, excluding prior history, and refit the forecast (Figure 6). The level and trend were about the same, but the lower forecast error cut the safety stock by 25 percent or $14,110.

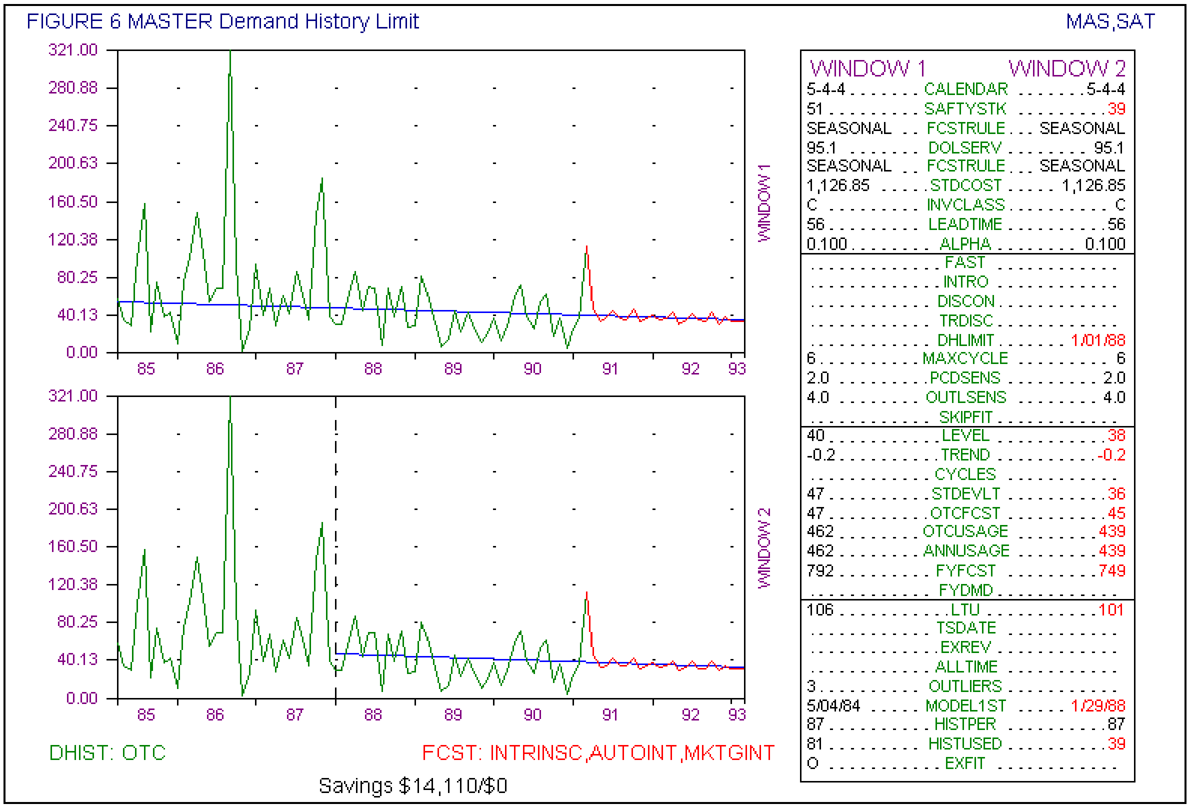


Figure 6 –Demand History Limit

Last month's demand was an unusual occurrence (initially flagged by the demand filter) and, upon investigation, we do not expect it to happen again. We tightened the outlier sensitivity (all FGS control limits can be set by SKU) for this item from 4 to 3 standard deviations and did another refit.

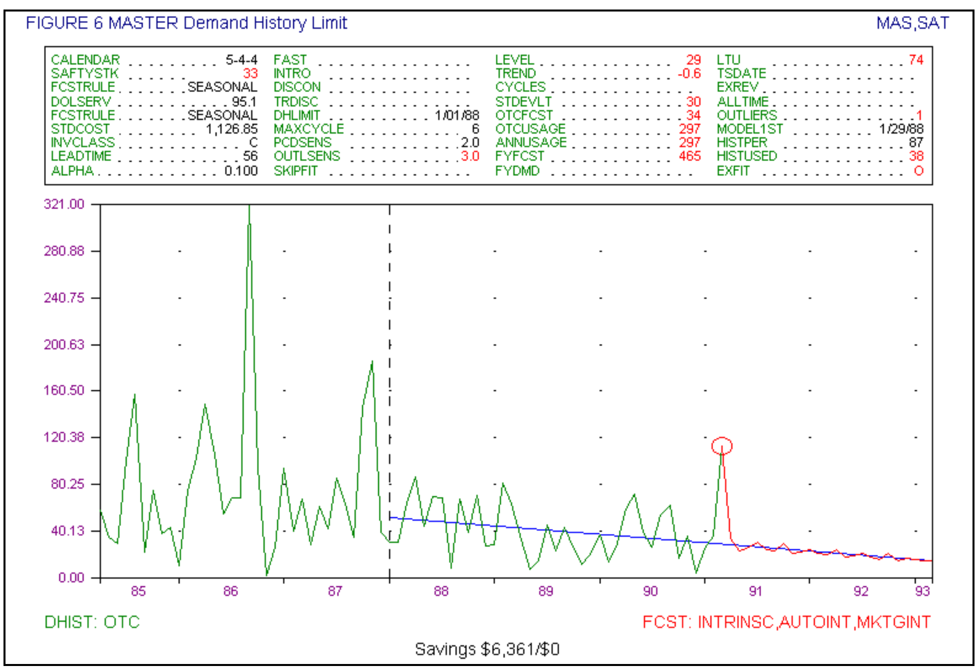


Figure 7 — Outlier

This caused last month's demand to be excluded. The forecast dropped from 45 (originally 47) to 34. Actual demand for the month came in at 34. The lower forecast error dropped the safety stock by another $6,361, for a total inventory savings over both changes of $20,471 or 35 percent. Not bad for a few minutes' work!

15. The Value Of Education

The product in Figure 9 showed up on the tracking list. We set the DH limit to January, 1990, to get rid of all the 1988 and 1989 volatility and the safety stock significantly improved. We filed the changes and were ready to go on to the next product when the inventory planner volunteered another alternative. She knew this product's market had been fundamentally the same since 1987. The only difference was some fire sales in 1988 to prevent shelf life problems, and some large promotions in 1990.

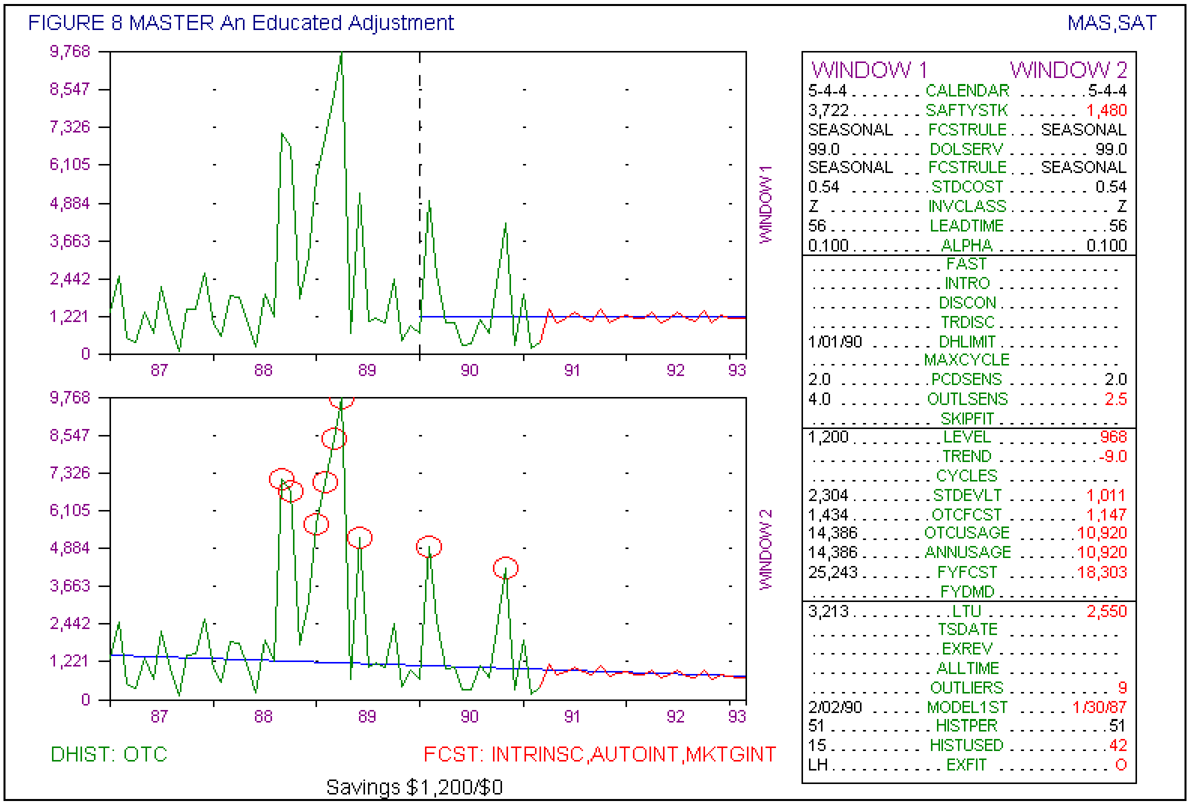


Figure 8 –An Educated Adjustment

We removed the DH limit (to use all the demand history) but tightened the outlier sensitivity from four standard deviations to 2.5 and refit. The system detected and ignored nine outliers generating a better forecast with a safety stock $1200 lower than the previous scenario. The forecast for the following month dropped from 1,434 units to 1,147, and the actual came in at 912.

Because she had just finished the APICS certification review in Master Planning, the planner had the knowledge and confidence to override the forecast. Clearly, GE Silicones achieved a quick return on investment for her APICS course.

16. The Payoff

At the end of every session with a planner, I print out FGS' Planned Inventory Report which shows the changes in safety stocks as a result of their work. For our April 1991 forecast review process this amounted to $87,000 in inventory savings for about eight hours of work. As a general rule, we save about $10,000 in inventory for each hour spent using the FGS Simulation Facility to review exceptions. This is time very well spent in my opinion.

From backups of the FGS database, I obtained the forecast errors from one year ago. I loaded the errors into a copy of the current database to simulate current conditions (lead times, costs, service levels, etc.). Then I calculated how much today's service would cost but using last year's forecast errors. The result, when extended to all our make-to-stock products is worth about one million dollars in safety stock reduction, with the same level of customer service. Besides this savings, there is the carrying cost on that amount which would be somewhere between $150,000 to $200,000 annually. Can we claim these benefits? We can if we consistently keep the forecast accurate. Part of the challenge is to keep the forecasts improving. If we do not continue to "work the forecast exceptions," the safety stocks will return to their old levels. If we can keep the forecasts improving, however, we can realize similar benefits every year.

The inventory planners now have confidence in their plans. They don't feel the need to fudge the plan to make up for a lack of confidence in the forecast. Rather, if they lack confidence in a forecast, they fix the forecast. Even the marketing product managers have more confidence in the forecasts; and when they raise a concern, it goes into the forecast as marketing intelligence. It either improves the forecast or the marketing people find out why it didn't.

Finally, because manufacturing isn't spending so much time producing the wrong products (because of erroneous forecasts), there is more capacity available to produce the right products.

17. Final Results & Summary

During the first year of the FGS forecast implementation, customer service (shipment to request) for our business went up 40 percentage points, while total finished goods inventory expressed as an average time supply remained unchanged at 3.7 weeks. The forecasting process is partially responsible. Some other important influences were:

1. Inventory planners assumed responsibility for their plan (and forecast).
2. Formal sales and operations planning was begun, and supply and demand are now fitting.
3. Our total quality process is resolving supply issues.
4. We are investing in de-bottlenecking the right processes.

The title of this presentation includes the word "forecast" which we all know means "watch out" and "throw out". This presentation shows the need for exception reporting, graphical visual review, and simulation to get the best forecast possible. We regard our forecasts as superior. Superior to what? To yours? No. The goodness of the forecast relates to the product and its market. We can, however, state that this month's forecast is superior to last month's which was superior to the prior month's. This is what continuous improvement is all about.

18. References

1. Greene (ed.), Production and Inventory Control Handbook (2nd ed.), APICS, 1987, p. 29.15.

19. About the Authors

Robert M. Duncan is principal at Duncan Consulting in Ballston Lake, New York (518-527-4503). As of this writing he was a Senior Analyst in the Materials and Logistics Department of GE Silicone Products in Waterford, New York. He was responsible for forecasting and inventory planning, sales and operations planning, and resource planning. Bob was an MRP project manager at Silicones and at King and Prince Seafood. He has also worked in scheduling, facilities planning, and industrial and process engineering. He served for three years as a Peace Corps volunteer in Columbia, South America, consulting in management and teaching college. He holds a B.S. degree in Industrial Engineering and Operations Research from Virginia Polytechnic Institute and State University. He is an active and certified (fellow) member of APICS.

John A. Estep is president of E/Step Software Inc., of Yakima, Washington (509-853-5000), a firm specializing in the development of software for manufacturers and distributors. With over 40 years' experience in education, consulting and systems development, Mr. Estep's expertise is concentrated in the areas of database design, demand forecasting, inventory planning, distribution, master scheduling, and quality control. In addition, Mr. Estep has taught courses for APICS certification examinations and has presented to APICS national, regional, special interest group, and chapter meetings. Prior to launching his own business, Mr. Estep worked in consulting, manufacturing systems development, and operations research. An APICS Certified Fellow, Mr. Estep holds an M.S. degree in Electrical and Computer Engineering from the University of Massachusetts, and a B.S. in Mathematics from Allegheny College.

The above is a reprint of an article submitted by Robert M Duncan and John A. Estep which appeared in the APICS 34th International Conference (1991, Seattle, Washington) Proceedings. Several charts and the accompanying text, which were deleted from the published article for production reasons, have been restored in this reprint.

For additional information, contact:

E/Step Software Inc.
12015 Summitview Road
Yakima, WA 98908

Tel. (509) 853-5000
Email: John.Estep@EstepSoftware.com
Web site: www.EstepSoftware.com