

Hed:How to Select A Dedicated Forecasting Software

By: David P. Reilly

Before you start searching for the right software, determine your forecasting needs ... forecasting software should accommodate your forecasting process and not the other way around ... forecaster must have a working knowledge of forecasting models, diagnostic tools, and error metrics.

David P. Reilly

Mr. Reilly is Senior Vice President of Automatic Forecasting Systems, Inc. He has distinguished career as an innovator and a developer of creative time-series applications. He is a former GE executive and Senior Manager at the American Stock Exchange. He is also on the Drexel Graduate Faculty. He has been a senior consultant to companies including Pepsi-Cola, General Mills and Anheuser-Busch.

While many forecasters use spreadsheet applications, many have found the benefits of using dedicated forecasting software (DFS). The search for such forecasting software should result in a solution that best suits a particular company's needs, circumstances, and expectations about the future. Simple forecasting problems lend themselves to less formal approaches and to simpler methods. Complex forecasting problems require more organized forecasting processes and statistical methods. Furthermore, they necessitate the use of more sophisticated tools and more complex solutions. DFS is an essential tool for a forecaster. The basic question is how one should search for and evaluate such a forecast software solution?

FORECASTING NEEDS

A dedicated forecasting software solution can support a company's need for data collection, analysis, model development and evaluation, and forecasting. It is important that whatever the software you select, it should be consistent with the specific needs of the company. In other words, the company's needs should not be driven by the software's needs. Therefore, it is important for a company to determine what its business needs are before looking for a software.

What does a company need to know before embarking on a search for a DFS solution? It is often helpful, but not mandatory, for a company to develop its forecasting process. The process should reflect the company's business processes, business decisions, and management information needs. A complete understanding of the nature of the variables to be projected, time horizons for forecasting, historical patterns experienced, and drivers of the actual outcomes is important. The levels of detail required in the forecast along with the data and forecast segmentations must be known before searching for a software solution. The types of management reports and forecasting presentations are also important to know in advance of software evaluation and selection. One can ask relevant questions about the DFS package that is under consideration only if it is known what one is trying to accomplish, the nature of the present and expected forecasting problems to be addressed, and the goals of the forecasting efforts.

There are many forecasting software solutions available, each with different features, different statistical models, and different levels of vendor's support. There is a wide range of performance features of forecasting packages but DFS offers substantially

more advanced characteristics when compared to spreadsheet applications and spreadsheet add-ins. Therefore, a systematic inquiry and search is important to find the best DFS solution for your needs — both for the present and for the future. Different companies will make a different selection based on their needs, but the budget allotted for DFS will also play a role in the selection.

FORECASTING SOFTWARE CONSIDERATIONS

There are two broad levels of consideration when evaluating a software package: (a) what does the package do and (b) how does it do it. The “what” provides insight into the potential of the package to solve the forecasting problems that are central to the business and to its management. The “how” provides insight into the effectiveness of the software in implementing features the company is concerned about. The “how” is just as important as (sometimes more important than) the “what” when it comes to forecasting packages. Hence, in developing questions for a software search, each “what” question should also have an accompanying “how” question to ensure that the software fits well into the company’s needs.

So, what are the key features a company should be concerned about when considering a forecasting package? The general concerns typically include:

- Expert System and Algorithms
- Data Preparation and Analysis
- Method Selection
- Model Diagnostics
- Aggregation and Segmentation
- Reporting Capability

Expert System and Algorithms

The expert system within DFS is essential for data analysis, as well as for forecasters' ability to produce quality forecasts. Most forecasters are not statisticians, nor do they need to be. Most forecasters deal with problems that require a huge staff of trained statisticians. But with a proper expert system, which is built into a forecasting package, the forecaster can achieve efficiency and quality of forecasts with minimal manpower. Hence, a clear understanding of the level of support that the software provides to a forecaster is essential. While the forecaster should be able to interpret and evaluate the information provided by the expert system, it should not be necessary for the forecaster to undertake and complete all of the functions of a statistician. Much has been systematized through the algorithms built into the dedicated software forecasting packages. The expert system capability is often described as the automatic forecasting system that automatically selects the best model for a data set to prepare forecasts. Discussing the expert system's capability with the software vendor and understanding its functionalities is an important part of the inquiry process.

Data Preparation and Analysis

Data preparation and analysis are a central consideration in the ability of a system to satisfy the forecaster's needs. A feature which is typically important is data cleaning along with stability testing. This calls for the use of a variety of statistical measures of central tendency and dispersion. Evaluation of the data for trend, seasonal patterns, outliers, shifts, and event-driven conditions is also an important feature of a software package. The presence of data transformation functions that can be applied to the data is important to the analysis capability of the package. Certainly, for analysis and

presentation purposes, the ability of the DFS solution to plot data is fundamental to the forecasters' mission. Visual graphic presentation along with thorough data analysis is fundamental to a satisfactory software solution for forecasting.

Method Selection

Statistical method selection is a most critical consideration in evaluating a forecasting package. The ability of DFS to use the data preparation and analysis and to match the statistical method to the data is extremely important. Since the issue is one of matching the statistical method to the forecasting problem and to the data pattern, the range of forecasting methods resident within a forecasting package is a prime factor in deciding on a software package. The ability of the package to build models that reflect the data history and the business factors are very important. So, in many cases, one may be looking for combinations of model types to produce the best forecast for management. A range of univariate methods (time series), multivariate methods (cause and effect), and ARIMA/Transfer Function methods is generally a desirable feature in a forecasting package. With that, the forecaster can use both "single-threaded" solutions as well as "multi-threaded" solutions for business forecasting.

The DFS solution should allow different orders of SARIMA (p,d,q) (P,D,Q) models with different lags and leads of exogenous variables either specified by the user or internally developed. The DFS solution should serve as a good proxy for a skilled statistician. Tests for changes in historical patterns should be done to ensure that it reflects only those historical patterns in the forecasts that have been shown to have a consistent presence.

Many developers of DFS, unable to duplicate the correct statistical analytical process of decomposing the observed set to signal and noise, limit their selection to pre-programmed alternatives (much akin to being forced to buy a suit off the rack without the benefit of tailoring). In an ideal DFS, historical data is partitioned into “training data” and “validation data” (or out-of-sample data), and alternative models are estimated using the training data and evaluated using the validation data, normally from a single origin. In many DFS, no tests of significance are ever performed, no tests of the Gaussian Assumptions are made, and no residual diagnostic checking is performed in the “fitting process.” To make things worse, if you change the number of “out-of-sample” values, the best model changes, which suggests the tail is wagging the dog.

Model Diagnostics

The approach of the software to model development (specification, estimation, validation, and forecasting) is a major area of concern in evaluating software packages. From a statistical perspective, it is best if the software can build a model that provides the highest degree of reliability and accounts for shifts, pulses, and causal factors. Level shifts in the data reflect the impact of such things as changes in the law and competitors going out of business. The ideal DFS solution should be able to identify automatically such level shifts and incorporate their effect into the model. Also, changes in trend need to be identified and incorporated. Unusual values also need to be identified and accommodated with pulse variables. Finally, seasonal pulses (the impact of Christmas, for example) should also be accounted for.

If the DFS solution does not perform these tasks, then it is left to the forecaster to do them, shifting a great deal of the “virtual statistician’s” work to the forecaster. It is important that the software package facilitates the consideration of alternative model specifications and their estimates. The presence of effective expert system algorithms is important to the ability of software to evaluate data and alternative model specifications and estimates for recommendation to a forecaster. (Of course, the ability to override these recommendations is also important to the forecaster.) Also, it is important that the software package gives a range of error criteria that the forecaster can use to select the optimal model for forecasting. Typically, one will select the model that produces the lowest expected error and/or lowest range errors. However, the final model selected must be scrutinized both for statistical significance of the estimated parameters (necessity) and for sufficiency of the model (i.e., the errors from a selected model are random.)

To determine the appropriate model for forecasting, the forecaster has to test a model with all the diagnostics tools at his or her command. So it is important that the software package has adequate test statistics, including R-square, t-test, F-test, Durbin-Watson statistic, and so on. The forecaster needs to have a working knowledge of the test statistics that he or she uses, and ensure that these are available within the software. The ability of the DFS solution to graph the actual data, forecasts and forecast errors is important for evaluating and presenting forecasts. Presenting a forecast range both numerically and graphically is equally important. Capturing assumptions and tracking their validity is also important to the forecast process and forecast analysis. The ability of the software to support tracking, analyzing, and remediating error is a key consideration in selecting the best forecasting package.

Aggregation and Segmentation

Different users of forecasts have different needs for aggregation and segmentation of forecasts and actual data. Typically the forecaster wants that the software package has the capability to slice and dice the data — both historical and forecast. So being able to look at the data and forecasts by hierarchy structure is important. These hierarchies may include product hierarchies, distribution hierarchies, geographic hierarchies, management structure hierarchies, customer configurations, and business unit hierarchies. With that feature, one can look at the data from different perspectives.

The forecaster is looking for a software support tool that has the necessary characteristics to produce adaptable, accurate, and meaningful projections for management. The ability of the software to minimize the amount of unnecessary data handling and manual data manipulation as well as unnecessary export of information for use by management is essential to the goals of a forecaster. Ultimately, the forecaster is a business analyst providing forward-looking information to management. He or she needs a tool that optimizes the use of time and professional energy and facilitates the analytical and communication capabilities.

Reporting Capability

Another important feature is the reporting capability of a package. Since each end-user has a different need of forecasts, they have to be reported differently. Some want the report in units and others in dollars. Some want them by category/brand, while others at an SKU level. Also, forecasts may have to be broken down by customer, channel of distribution, region, and country. The software should be able to do all of these.

QUESTIONS TO BE CONSIDERED

The search for a DFS solution is a systematic inquiry into the needs of the company, as well as into the capabilities of the software. In order to get the best results, it is best to have an agreed upon set of questions to ask vendors that offer a software solution. The questions are intended to ensure that the company gets a complete solution to its business and forecasting needs. Below are questions that can be used as the starting point of inquiry when searching for a software solution. These can be of great help not only to those participating in the forecasting process but also to the IT personnel who are assisting in the search process.

Data Preparation and Analysis Questions

The heart of a dedicated forecasting package is the data preparation and analysis. All other forecasting-related activities are contingent upon the data, data pattern, and business variables being forecasted. So, which questions should be asked by the forecaster? Below are some of the most critical questions.

1. Which means of data collection and data loading are available?
2. How are data accessed and how easy or difficult to access?
3. How does the software segment (slice and dice) the data?
4. Does the software decompose data into separate elements — trend, seasonality, and cyclicity?
5. What type of data plotting capability does the software has? How flexible and user friendly is this capability?
6. What types of data transformations are possible (e.g., logarithms, inverse transformation, power functions, etc.)?
7. How does the software identify and handle seasonality and cyclicity structure in the data?

8. How does the software identify and deal with missing values?
9. How does the software identify and deal with shifts in the data?
10. How does the software identify and deal with outliers in the data?
11. How does the software identify and deal with distinct time trends in the data?
12. How does the software handle lead and lag structures?
13. Which metrics (such as cross-correlations, tests for constancy of variance, and constancy of the parameters) are produced by the software in undertaking the data analysis?
14. How does the software “verbalize” the model and its coefficients so that different members of the forecasting team are comfortable with and can readily understand and buy into the forecasts?

Note: Be sure to add a “how” question to each “what” question posed to the vendor.

Method Selection Questions

Data analysis is fundamental to the selection of the right model. The model is designed to reflect the data pattern (i.e., the “memory” within the history) as well as the business factors that may be driving the forecast. Time series and ARIMA models deal with the “memory” element, while regression models, transfer functions, and econometric models deal with the cause-and-effect elements. (A transfer function is simply an optimal combination of ARIMA models supplemented with cause-and-effect elements.) Hence, it is important that the software is capable of supporting both time-series and cause-and-effect models, as well as capable of using them in combination. So, which questions should be asked by the forecaster? Below are some of the most critical questions on this subject.

1. Which models and statistical methods are resident within the DFS solution? How are these models evaluated and implemented by the expert systems? Does the DFS test the adequacy of a model? How does it “grow the model?” Does it incorporate either leads or lags of the variables or augmentation via an error structure?
2. Can the package produce a model that has never been produced before, i.e., can it tailor the model to the data vis-à-vis forcing the data into a pre-programmed set of models?
3. Which criteria are available for model selection purposes? Can multiple criteria be used for this purpose?
4. How does the package fit alternative models to the data set?
5. What level of expert system capability exists to qualify and select models for consideration?
6. What level of user override is available in the creation and selection of models for consideration?
7. Which test statistics are available for use in validating and evaluating alternative models and model specifications, e.g., statistical significance, autocorrelation, multicollinearity, and heteroscedacity?
8. How does the software deal with uncertainty, e.g., range forecasting and simulation?
9. How does the package recognize patterns within time series modeling?
10. How does the package deal with regression for developing and validating a model?
11. How does the package deal with seasonal factors, cyclical factors, event factors, leads and lags, data transformations, and other such considerations affecting model development?
12. Which metrics are included in the software to support model diagnostics?

13. Which metrics are included in the software to support error measurement and analysis?
14. How does the software handle “in-sample” forecasts?
15. How does the software handle “out-of-sample” forecasts?
16. What types of graphics are available for model development and model evaluation?

Note: Be sure to add a “how” question to each “what” question posed to the vendor.

Forecast Implementation Questions

There is a variety of concerns related to forecast implementation. Some of the questions that forecaster should ask are given below.

1. How are the updated actual data collected and loaded on to the forecasting package?
2. How is the forecast and resulting error updated, tracked, and analyzed within the forecasting package?
3. Is large-scale batch forecasting available within the software?
4. What levels of aggregation of forecasts can the package handle?
5. Can the forecast reports be compiled by segmentation type, e.g., by function, product and product category, and geographic territory?
6. How does the software handle “top-down” forecasting?
7. How does the software handle “bottom-up” forecasting?
8. Which management reports are available from the software?
9. Does the software report on the model used?
10. Which measures of error are available for analysis?
11. What are the “drill down” and query capabilities within the forecast and the actual data?
12. Can the software evaluate the data and create analytical statements, such as “.....the following list of SKUs experienced a significant level shift starting such and such periods ago?”
13. How are assumptions identified, quantified, and stored?

14. How are the forecasts and forecast assumptions and reports shared with other users?
15. How are the forecasts and forecast assumptions and reports shared with management?
16. What types of graphics are available for the forecast?

Note: Be sure to add a “how” question to each “what” question posed to the vendor.

Software Support and Other Requirements Questions

Assuming that the software has the desired combination of features and capabilities, the last set of considerations relate to the “mechanics” of the software.

Typical questions to be asked should include:

1. What are the query capabilities of the software?
2. What are the report writing capabilities of the software?
3. What is the presentation capability and format of the software?
4. What level of graphics and graphics presentation exists within the system?
5. Does the software have batch forecasting capabilities in addition to real-time forecasting capability?
6. What data bases and software interfaces are supported by the software?
7. What are the hardware requirements for the software?
8. Does the package interface with the company’s existing hardware and software platforms?
9. Can the software be readily linked to other data bases and software (internal and external)?
10. Can the software be readily interfaced with other dedicated forecasting packages?
11. Can the software be readily interfaced with specialized packages, e.g., neural networks and econometrics?

12. What is the frequency of updated software releases? How are these distributed? Are they cumulative in nature or are they independent releases?
13. What level of training is available for the software users? What is the quality of the manual and reference material for users?
14. What type of vendor support is provided on an ongoing basis?
15. What level of vendor support is provided during the installation and set-up of the software?
16. What level of consulting support is available and/or required in the installation and set-up phase of the software implementation?
17. Which other companies are using this software? What has been their experience? What is their degree of satisfaction?
18. Will the vendor provide a list of references of users of its software to determine their level of satisfaction with the software?
19. What are its limits in terms of data base size, number of variables, data length, forecast time horizon, and data export and data import?

Note: Be sure to add a "how" question to each "what" question posed to the vendor.

SUMMARY AND CONCLUSIONS

The set of considerations and questions presented in this article is not exhaustive, but it gives a good starting point. Supplementing the above with other questions pertinent to the specific situation and concerns facing the forecaster, the company, and its management will have a broad basis of inquiry to ensure a better software selection decision. The objective is to find the solution that best fits your company's needs and circumstances.

Ultimately, one is entering into a partnership with the vendor and its employees. Be sure to meet with the vendor's management, development staff, and support staff. Talk with them in order to develop a sense of their commitment and capability to deliver the services they promise. Make sure that the vendor is financially sound. This will minimize the risk of unexpected disruptions of service or withdrawal of vendor support.

The bases for selecting dedicated forecasting software are grounded in the specific needs of the company, its forecasting process, the business decisions to be made, and the application of appropriate data and statistical methods for forecasting. These requirements should be fully defined and explored before undertaking a software search. An independent consultant can sometimes be of great assistance in giving you an unbiased perspective in these regards. It is important to consider both current and future needs. Make sure that the selected DFS package has the flexibility and expandability that will be required for both its current and future needs.

Certainly a forecaster's education is important. Before embarking on this venture, make sure that the forecaster has a working knowledge of various metrics, analytical tools, and methods and models that can be applied to a forecasting problem. This way you will be sure that you are getting exactly what the company needs (not just what a particular software solution may offer). While it is not necessary to be a statistician to effectively use the available software with the embedded expert system, the forecaster must have a reasonable level of knowledge of when and how to use different methodologies in order to make informed judgments about the best approach and the trade-offs that are inevitable in such an undertaking. So a forecaster with a good working knowledge of forecasting methods as well as of the business and its operations is in an

excellent position to take the forecasts to a higher level of performance by selecting the best dedicated forecasting software solution for his or her company.

Properly structured, the search process can result in a software package that will significantly improve the ability to forecast as well as the management's decision process. It can also improve the productivity of all who are involved in the forecasting process of the company. The search is an exploration where asking the right questions and getting clear feedback from vendors are essential. Structure your inquiries, ask follow-up questions, check vendor references, and be sure to become familiar with all of the key personnel of the vendor. It is ultimately a business partnership, and your forecasting success depends on it.

Here is a last suggestion that can assist a great deal in selecting the best software solution. Create a test data set or sets that best characterize your company's forecasting problems. Share these with the potential vendors (holding out a sample to which their forecast results can be compared). Ask the potential vendors to develop the optimal model(s) and produce a forecast of the hold-out period, and do so within some designated time period. With that one can determine the effectiveness of the software. Since the vendor is most knowledgeable about the forecasting software it is offering, this exercise will provide "live" insight into the speed and ability of the software package in dealing with your forecasting problem. It also affords an opportunity to see how the software presents the forecast, does the data analysis, and verbalizes the model, key assumptions, and other forecast issues. These can then be compared across vendors to determine the best software solution for your company. Here, you may like to find an independent consultant or someone else you trust to help you in evaluating different software results.

A very powerful way of assessing different Dedicated Forecasting Solutions is to attend a forecasting conference like the ones conducted by the IBF. Bring your data with you in an MSEXcel file or a basic text file and walk around the booth area and ask several vendors for a live demonstration on your data. Some vendors only want to exhibit their software on their data, which oftentimes is trivial. Find out what their software will do with your data!

The key to finding the best software solution is to know your company's needs, know the forecasting models and metrics of most importance to your company's forecasting problems, and ask the right questions. Surprises after the fact usually do not come from the inquiries that were made, but rather from the questions that were not asked. So ask the right questions. ■