CARREKER	**Autobox: iCom® V2.0 Forecasting Engine**

Last Updated: October 1, 2003

Question:[1]

"Why has Carreker changed forecasting platforms from SAS® to Autobox®?"

Answer:

Answering this question requires two brief essays. (1) a historical synopsis Carreker's forecasting history; (2) a somewhat technical synopsis of Carreker's dissatisfaction with SAS as a cash-demand forecasting tool.

(1) HISTORY

Since entering the cash management business, Carreker has dedicated itself to the corollary task of developing the finest cash forecasting capability humanly possible.

This Quest began with Carreker's Cash Forecaster® that pushed the envelope of neural-net technology. Although promising, it became clear to us that neural-nets were technically limited with regard to the relatively high-frequency sort of daily data that real-world banks deal with. Given their high complexity, they performed no better than "traditional"[2] (best-model) forecasting methods that were faster and less resource intensive.[3]

ICom took forecasting to the next level allowing for modular forecasting engines to be "plugged in" and "out" of the application to ensure customers are always given the most accurate forecasting available. Hence, with the launch of iCom, a traditional method was adopted with SAS as the chosen statistical modeling tool. SAS consultants provided a turn-key forecasting program that replaced Cash Forecaster's neural-nets with no loss in forecasting accuracy and by some measures even a bit of improvement. However, being an "external" program to iCom, it requires independent installation, program execution, and maintenance separate from iCom's central system and installation files.[4]

Nevertheless, these are secondary issues since highest-possible forecast accuracy is, as we have stated, the supreme goal of our forecasting efforts. In this regard, SAS is problematic in certain areas of automated cash demand forecasting, particularly with respect to exogenous ("level") shifts in trend, missing and extreme value mitigation, and pulse (event) detection. Therefore, in early 2002, as a part of routine maintenance and development,

---

[1] Please note this document is essentially one of the "frequently asked questions" (FAQ) that will become part of the official "iCom V2.0 Cash Forecasting White Paper."

[2] "Traditional forecasting methods" generally refer to using a statistical regression tool (i.e. SAS, SPSS, Eviews, etc) to make a forecast by passing data through a finite set of previously *a priori* defined "models", then selecting the "best" model through some other *a priori* defined selection criterion.

[3] This experience has been confirmed in the forecasting literature. For instance, see: "Principles of Forecasting: A Handbook for Researchers and Practitioners", J. Scott Armstrong, ed. Kluwer Academic Publishers: Boston. (2001), pp. 245-256.

[4] This at times has caused concern for our client's IT and Security staff.

Carreker began a thorough and rigorous search for a new "next generation" forecasting engine.[5] In summary, the technical selection criteria was:

1. Ability to *automatically detect:* level-shifts (interventions), seasonal pulses (lead & lag events), and transfer-functions (trend volatility) needed to decode patterns from the "noise" of relatively dirty and volatile cash-demand data. In addition, automatic outlier & missing value mitigation is essential. When forecasts fail, they generally fall short due to these "usual suspects" that traditional statistical packages cannot deal with *on an automatic basis*.
2. *Internal (non-executable) progra*m that can be seamlessly integrated within iCom. Must possess a small IT "footprint" for both system resources as well as security purposes.
3. Speed. Ideally, an internal program where data and forecasts are passed back and forth *in memory* rather than disk.

(2) TECHNICAL

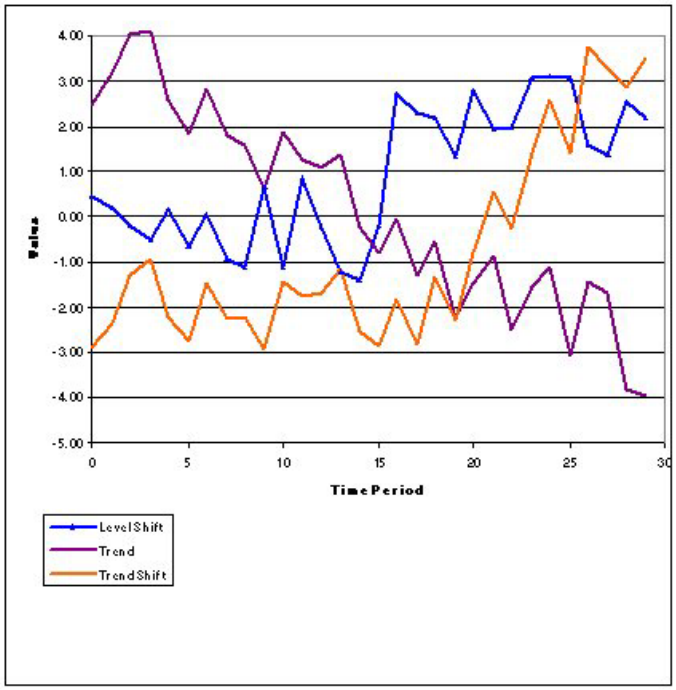
Definitions:

- **Level Shift**
  Significant and sudden change in trend (mean) on a moving average basis (see Graph A below).
- **Pulse**
  Extreme value relative to the standard deviation. Can take the form of one-shot (non-deterministic) outliers, or reoccurring seasonal (deterministic) patterns (see Graph A below).
- **Transfer Function**
  A mathematical statement that describes the "transferred" characteristics of a system, so that when applied to the input of the system, will fully describe its output. Much of the "art" of ARIMA forecasting involves the identification of the appropriate transfer function between input (historical data) and output (forecasted data).
- **Mitigation**
  The discovery of and numerical implementation of the above threats to internal forecast validity.

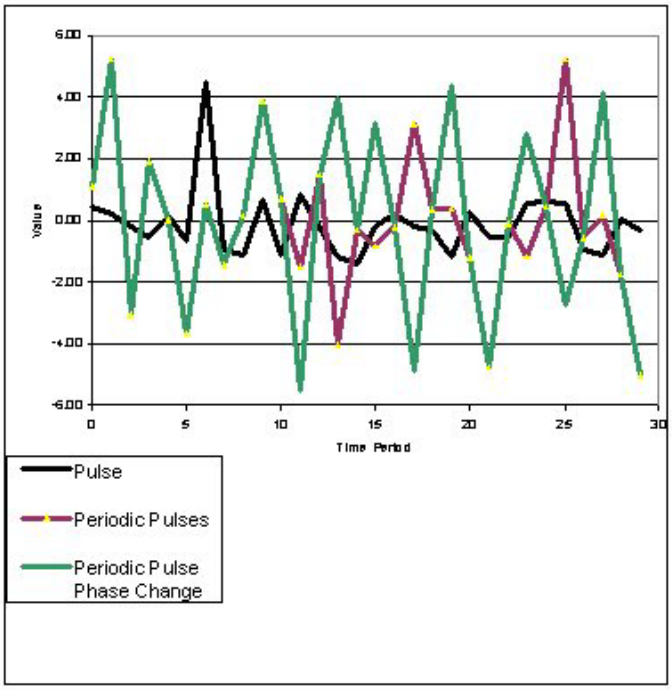
[5] We investigated literally every serious statistical package, data-mining system, applets library, and forecasting option on the market -- over 100 forecasting systems – and even seriously considered an "in house" algorithm. In the end, we settled on Autobox as nothing else even came close for our purposes. Others have independently come to the same conclusion. For instance, in the Tashman-Hoover academic study, Autobox was scientifically ranked "best" business automated forecasting package (J. Armstrong, 2001).

GRAPH A

Level Shifts



Pulses



Another way of looking at these threats relates to "what is the optimal amount of data to use in forecasting?" and/or "when does my data become obsolete?" Most so called data mining products do a very poor job answering the above questions and either weight too heavily or too lightly the past. For instance, if a new mall opens across the street from an ATM machine, the historical data becomes obsolete and using it will produce biased forecasts. As things change, as they do in real-life (as opposed to the laboratory), the statistical model must change immediately to compensate. Hence, too much data being considered is just as much of a problem as too little – both are serious threats to valid forecasts.

Can SAS properly detect, process, and mitigate these threats and be used to produce accurate forecasts?
Yes, we believe it *can*[6]

However, the statistical translation of our original question becomes:

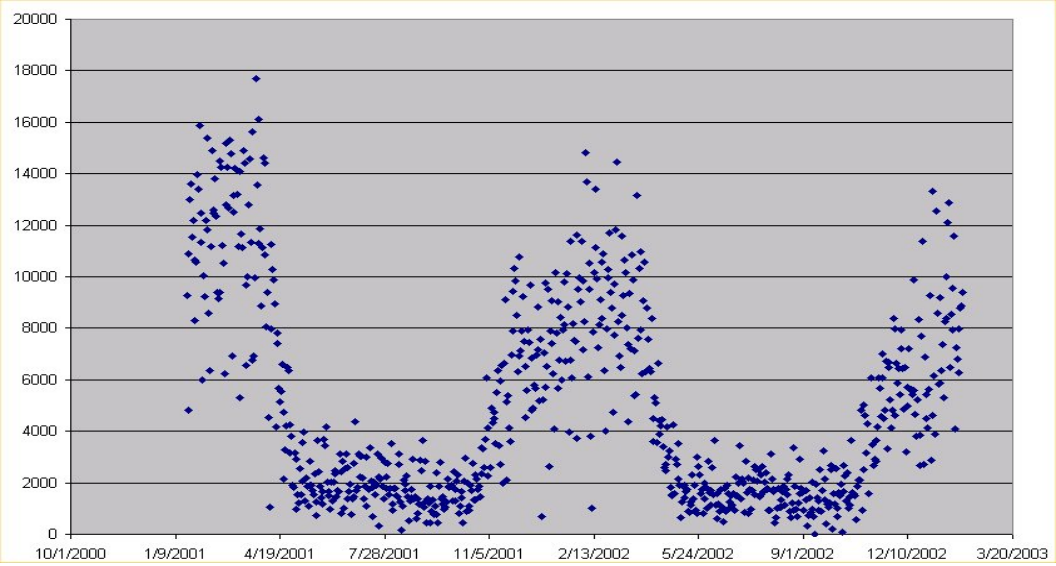
" Can SAS properly detect, process, and mitigate these threats and be used to produce accurate forecasts *on a purely automatic basis*?
No, we do not believe it can.[7]

[6] Given, of course, sufficient time invested by a well-trained econometrician personally at the helm of every forecast made. This would be very expensive and would require a small army of econometricians.

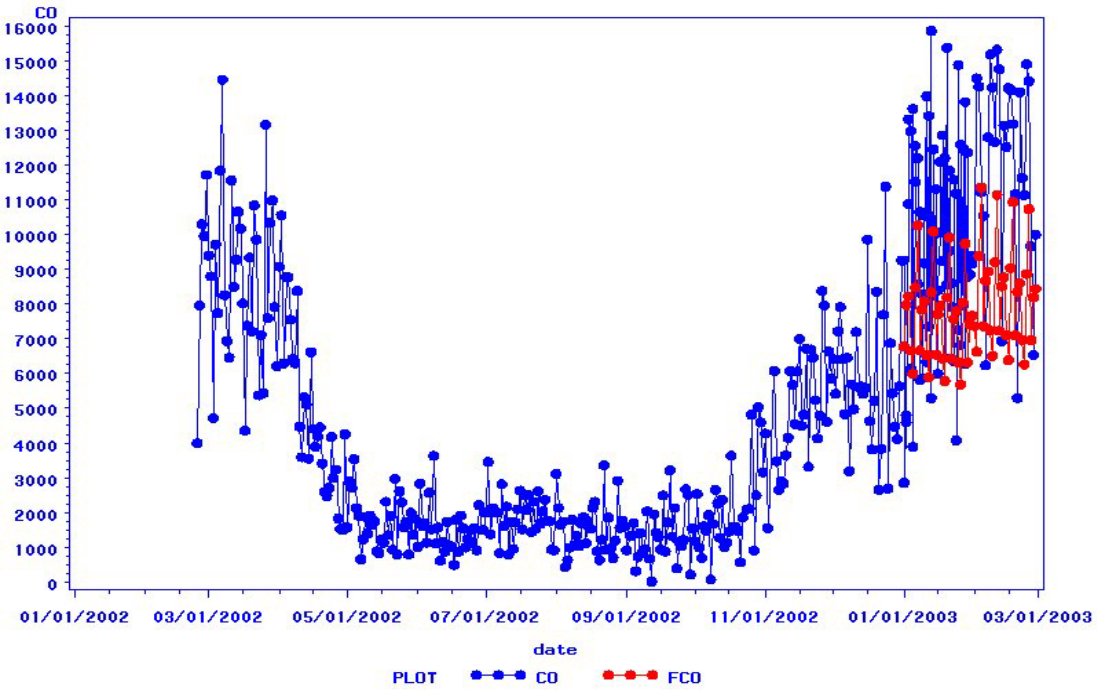
[7] *Nothing* in this document should be interpreted as disparaging of SAS as an organization or software product. SAS is an outstanding DBMS and Statistical tool that this author has used for years and will no doubt continue to utilize and enjoy. Within these numerous

Sometimes a picture truly does speak a thousand words. Hence, consider the graphs below:

Graph B: ATM Data - level shift(s), pulse(s), holidays, and outliers[8]



Graph C: Model-Centric forecast[9]

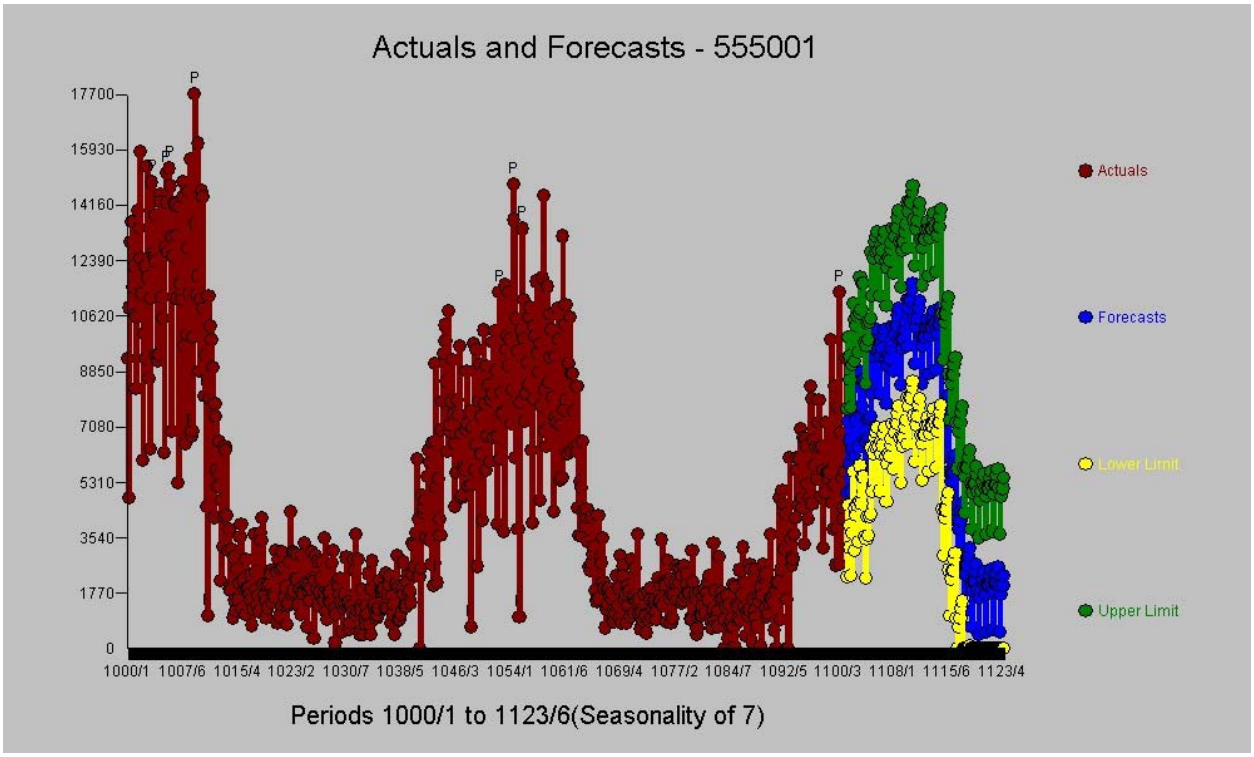


---

strengths, few of its competitors come close in this author's opinion. Indeed, the "due-diligence" testing analysis of SAS vs. Autobox has been done using SAS software! Unfortunately, *automatic* forecasting is not one of them. As a skilled econometrician, the author does not require an automatic forecasting and/or data analysis system – however, our clients do, hence their interest in iCom. Essentially, this document is referring to any "model-centric" statistical package, and is not specifically referring to SAS itself. In this regard, please see footnote 2.

[8] Graph (Excel) of an actual ATM machine's cash demand. This series is illustrative as it contains every "threat" this document has addressed.

Graph D: Autobox Forecast[10]



The above ATM depicted in Graphs B-D above, is "difficult" from a forecasting perspective. Even so, Graph D clearly shows that Autobox was able to mitigate all the "threats" we've spoken of and has produced not only an accurate forecast, but a seasonally accurate one as well.

Furthermore, further testing across 100 randomized ATM's show Autobox to enjoy enhanced forecast accuracy of 10-500% above our current SAS system. On an average total basis, on these ATM's (Florida), forecast error was reduced almost 50%![11]

Autobox is the primary software product from Automatic Forecasting Systems (AFS©)[12], whose main principle, David Reilly, was one of the early pioneers of the advanced time series modeling (Box-Jenkins/ARIMA) used today by academics and professional econometricians around the world. Hence, AFS brings to Carreker clients more than thirty-years of forecasting experience embed within the internal algorithms of Autobox.

---

[9] Graph (SAS) of actuals vs. forecasts using code provided by SAS consultants.
[10] Graph (Autobox) of actuals vs. forecasts from "full automatic mode" (no human interactivity).
[11] This was only a partial test (100 ATM's) without the scientific rigor necessary for a bona fide "metric". Nevertheless, ongoing scientific test results from more than 5000 randomly selected entities across branches, vaults, & ATM's have been similarly impressive. Once compiled, exact scientific results of this analysis will be appended to the final release version of this document.
[12] Visit autobox.com for more information.

As an expert system, Autobox *automatically* models and forecasts both univariate and multivariate time series via a single equation incorporating either pre-identified causal series or empirically identified dummy series, as discovered and statistically significant. Outliers or intervention variables are automatically detected and incorporated into both the non-causal (univariate) and the causal (transfer function) data structures. The set of pre-identified series can then be either stochastic or deterministic in form upon which proprietary algorithms will search for the most appropriate model given the "truth" of the empirical data.

Hence, the data suggests the model (not the other way around), then it proceeds to statistically evaluate numerous possible models and more importantly, parameters, that have been suggested by the data itself. In practice, a realistic limit is set on the maximum number of model form iterations (since the theoretical maximum is infinity). However, the exact specifics of each tentative model is *not pre-set*, allowing the forecasting power of Autobox to emerge. Indeed, the kind and form of the tentative models may have never before been tried, since it is the data itself that "speaks" thereby suggesting the necessary and sufficient conditions of the iterative process.

Autobox exploits both ARIMA[13] (univariate) modeling and Transfer Function (multivariate) modeling, most importantly allowing for the automatic inclusion and testing of exogenous interventions, need for transformations, and the addition or deletion of model parameters. Correction of autocorrelation (both traditional and robust), partial autocorrelation and cross-correlated residuals, and tests of significance are calculated as needed.

One of the most powerful features of Autobox is the inclusion of automatic intervention detection capabilities in both ARIMA and transfer function models. Almost all forecasting software *allows* for interventions to be manually included in a regression model, however they generally do not address how sensitive all forecasting methodologies are to the impact of interventions, missing variables, and/or the influence of exogenous changes outside the current model. If a data series is impacted by changes in the underlying process at discrete points in time, both ARIMA models and transfer function models will produce poor results. For example, if a new shopping mall opens near an ATM, this exogenously changes the level of demand for cash ant that particular machine. Without a control variable to account for this change the forecast model will perform very poorly. Autobox implements ground breaking techniques which quickly and

[13] "Auto-Regressive, Integrated, Moving-Average" that is the pinnacle of time series forecasting. Non-technically speaking, essentially it is a scientific art form that weighs varying observations with varying weights, thereby creating a premium or obsolescence within the history of the data.

accurately identify potential interventions (level shifts, seasonal pulses, single point outliers, and changes in the variance of the series).
In conclusion, being an excellent statistical analysis tool in the academic sense is insufficient. Rather, an *automated* data mining system must be able to, well …, *automate* human thinking, not merely select from a collection of a priori defined "best" models. Data should not be "thrown" at models, rather models should be employed based upon the data. Autobox *heuristically* (as opposed to ad hoc "best model" selection) analyses a particular data series then assigns the appropriate control variables necessary to create the proper transfer function. The result is considerably more accurate forecasts that actually approach those made on a custom basis by a skilled and experienced econometrician.[14]

Hence, with pride and confidence we will utilize Autobox[15] in iCom 2.0 and above releases.[16]

[14] For instance, the author is a trained and experienced econometrician and can always "beat" Autobox *on a particular series* given sufficient time. However, Autobox takes mere seconds to make the forecast and the author takes several hours of identification, testing, measuring, specification, and experimenting -- an intellectual and not to mention artistic process that Autobox comes very close to perfectly mimicking within its heuristics, and at a fraction of the author's salary.

[15] Carreker is absolutely committed to ensuring iCom offers the most advanced and best forecasting currently possible within the cash demand-management space. Therefore, iCom development philosophy has always been one of "modularity," meaning that we purposely write our software to flexibly accommodate *any* forecasting engine. Hence, we can change engines with ease relative to most of our competitors who are too often *structurally* tied to a particular statistical platform.

[16] Although we are thrilled to be partnering with AFS, especially considering their continuing commitment towards innovation, if and when anything significantly better emerges we will adopt it. AFS knows this and our clients should too.