In This Issue

The Missing Analysis Step Novices Ignore: Pre-conditioning Input

You have a great idea that you believe will consistently return profits from trading the market. You are confident that success will be yours because you have tested your idea with clean CSI data. You have carefully refined parameter settings and you have received approval from your friends. You are now ready to try your plan.

If it performs well in hindsight, will you be comfortable investing the store? If you lose on the first few trades, will you give up? Do you know in advance how much tolerance you will have for losses? Will you be tempted to intervene and go against your well researched result when times are good? Or bad? Will it disturb you to see great wealth accumulate in your account then abruptly disappear? Will you have the courage to continue if your initial experience reduced your assets by 25 to 50 percent? How good will you be in coping with an injured ego?

These are some anxiety provoking questions you will likely have as you proceed down the path of investing in your idea. You better address them now and dispel as many of the pitfalls and misconceptions as possible before embarking upon your trading adventure. Many doubts and questions can be avoided if your underlying notion is carefully and prudently tested before implementation.
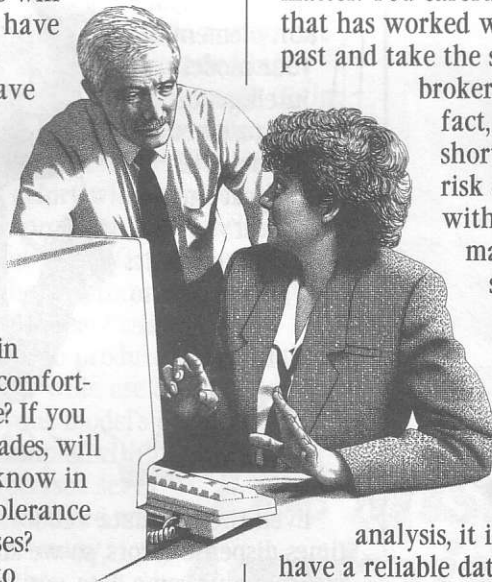
Most beginner analysts believe that trading the markets with one's own systematic procedure is a simple matter. You carefully design a model that has worked well in the simulated past and take the signals to your broker, right? This may, in fact, be possible from a short term view. At the risk of dousing your idea with a little cold water, may I respectfully suggest that some critical steps may have been left out.

The Importance of Accurate Data

Before doing any type of analysis, it is critical that you have a reliable data source for clean, accurate information on a multitude of pertinent markets which support your model's financial area of interest. Don't patronize a vendor who cannot provide backup source material that will authenticate suspicious price and volume reports. Avoid newcomer vendors who may have borrowed their data and are likely incapable of answering questions about it. If you doubt your data set, you can never be absolutely confident of your model's capability.

If your data comes from a questionable source, examine the data and verify its accuracy. Do not attempt to utilize a time series that is corrupted with zeros or spikes.

It isn't unlikely for a model, operat-

CSI introduces data on CD-ROM at our lowest prices ever. See enclosed flyer for details.

ing in the training (optimization) mode, to focus on data errors when developing variable coefficients and parameter weighting. I have seen simple errors in misstated highs or lows develop highly profitable, but flawed simulated trades. The sales, made in hindsight, were made at the peaks and purchases were executed at the lows, which, unfortunately, did not exist in the real market. Your data set may be the most important element of your model's intelligence because your model will expect the observed characteristics of the past to repeat in the future.
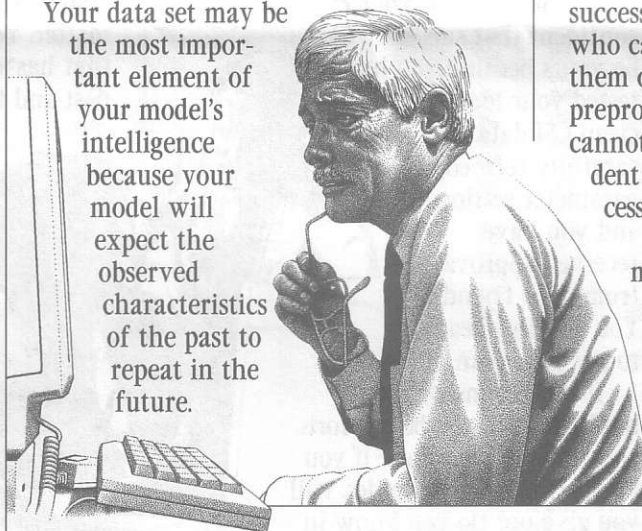
Even the best data vendors sometimes dispense errors, so we suggest you introduc some data verification control into your model. If you rely heavily upon certain critical input on a day-to-day basis, such critical input should be verified with another independent series. For example, if your model requires cash Deutsche-mark as an input, check to see that the direction of the nearest D-Mark futures corresponds with your cash data. Give yourself an error alert warning when data is suspect. Thinking ahead to such simple, but critical issues may save you much pain later.

The Importance of Preprocessing Input

The next important step is to condition the data through preprocessing filters and develop indicators from them to simplify and enhance your model's performance. Only after acquiring clean data and preconditioning that data should you build, train and test your model.

Many market students have readily rejected the idea of using neural networks to predict the markets because of poor performance. While neural networks may not be the best approach for some model builders, properly designed neural networks are often successful for many analysts. Those who cannot achieve success with them often leave out the critical data preprocessing step. Neural nets simply cannot do a good job without independent input data that has been preprocessed into a generic or relative form.

To give a neural network raw market data as input is akin to speaking Russian to an Englishman. Neither would know what the other is trying to say. Neural nets cannot make sense out of raw data input. They do much better with many kinds of information that have been preprocessed into a dimensionless statistical form. A good preprocessing effort will provide your model (neural network or other) with the intelligence needed to launch inputs on an equitable basis with similar rules. Your model can then concentrate on weighting inputs according to their respective value and influence on measuring profit performance.

It has often been said that the effort to produce a neural network is 90% input preparation. Ten percent falls into the category of input analysis and the measurement of the output response. Likewise, multiple variable input problems such as areas of intermarket analysis require a similar distribution of effort. A given market's movement is often influenced by other markets that may appear to be unrelated. All markets coexist. None operate in a vacuum.

Restructuring Market Input

Preprocessing is the step that accomplishes much of the necessary statistical analysis that removes the actual dollars and cents from the data. An example of a time series that might be considered as a basic input comes from an examination of crude oil. The analyst should develop a series that would represent the standard deviation of the accelerated movement of crude oil prices over time. This would produce a series that shows the relative movement of crude prices with respect to the recent or distant past. Doing a good job here might require that you know something about the leading and lagging relationships of crude on the value of the dollar. Similar relationships with recent yields in treasury bonds might be studied.

Determining Input

If it isn't clear what the preprocessing step entails, perhaps a couple more thoughts may get you pointed in the appropriate direction. The preprocessing step is where you identify pertinent input. Some of the focus involves reducing the number of input variables from a proposed list of candidate control parameters. It involves selecting the best forms and types of filters. Don't forget Kalman. It involves the weighting of parameters. It involves recognizing correlated market series and the introduction of fundamental indicators. It involves knowing how to introduce leading and lagging relationships that explain cyclical influences and forces on the market you wish to predict. And it involves examining how the performance of related financial market sectors influence your particular problem.

If you are attempting to solve a commodity market problem, don't ignore the financial markets as possible information sources. Solving a problem that involves determining the price of gold might easily require knowing about a host of related markets. Other precious metals, the gold mining stock sector, alternative investments in Treasury Bonds, The Fed Funds rate, the Federal Discount rate, the money supply and other fiscal and monetarist statistics might all be involved.

You really don't need to know too much about interrelated markets. After all, it is your model that will sort out the good and pertinent from the bad and irrelevant. Your model operates upon the data in a way that combines all the preprocessed input into a workable prediction tool. Giving adequate attention to the preprocessing step will determine how your modeling effort will approach the prediction problem you identified. A good model will efficiently sift through the preprocessed information set and will create a much better opportunity to produce a substantive result. Make wide use of ratios to enhance your model's performance.

For example, within the preprocessed set of information there are likely to be many series that are correlated. Taking ratios of correlated series helps to give your model the intelligence to know when one is getting significantly out of step with the other. Look at market pairs like crude oil and gold, crude oil and cotton, crude oil and aluminum, gold and silver, corn and soybeans, soybeans and soybean products, T. Bills and T. Bonds, Swiss Franc and the Deutsche Mark, etc.

Dealing With Correlated Input

Using correlated series as separate inputs in the same model could cause trouble in the training stage of a

The Missing Analysis Step Novices Ignore: Pre-conditioning Input

Ask Customer Service

Q. *I enjoy retrieving large amounts of history on-demand, which I keep in just the CSI format. I have left the Default Path blank in User Constants because I don't use the MetaStock format. During history distribution, QuickTrieve® beeps and displays an error message saying there is no place in the CT/MS list for my data files. How can I eliminate this unnecessary message?*

A. This error message occurs when QuickTrieve tries to create MetaStock-format files but there is no designated place to store them. The solution lies in User Constants, but not in the Default Path entry. Page 5 of User Constants includes two entries that are involved in Metastock-format file creation. We suspect that your entries for Create MS Daily and Create MS Hist are (Y)es. Change them to (N)o if you don't ever want Metastock-format files. This should eliminate the beeps and the error messages.

Q. *I'm having a hard time getting QuickTrieve to create back-adjusted continuous contract files. Gaps appear at the rollovers, which seem to result from an inconsistency in rollover dates. How can I get the rollovers to run smoothly?*

A. The most common problem with continuous contracts occurs when just one contract is available to the system for any given day. In this scenario, the rollover dates must be synchronized to make the necessary contract available to the system at rollover. This is sometimes easier said than done. When an automatic rollforward portfolio has been requested from CSI, the "rollover date" is the first day the new contract will be delivered. In the continuous contract algorithm, the "rollover date" is the last date a given contract is included. To get this to match up, you'll need to make your CSI rollover date one day later than the rollover date listed in the Continuous Contract menu's Change Program Options section.

A simpler solution is to maintain two or more contracts for each commodity at all times. This usually eliminates the problem of required data being unavailable.

Q. *I get data from CSI, but store it only in the Metastock format. Can I use QuickManager's® continuous contract feature?*

A. Technically, you can. However, from a practical standpoint, it is likely more work than you'll want to do on a daily basis. It requires that you maintain all the source data files in CSI format and an additional continuous contract file in CSI format. The continuous contract files can then be converted to Metastock format when needed. The MetaStock-format continuous files will only be updated when you add more data through the conversion routine.

Q. *Can CSI Customer Service help me with analysis using a third-party software program?*

A. We are happy to help you install your QuickTrieve and get your data files in order for analysis with your other programs. Many vendors have provided us with preferred portfolio information and details on required path names etc. If we have this information for your program, we'll certainly share it with you. You can help us with this process by having your third party software manual on hand. We'll support you in all aspects of data retrieval and maintenance, but will refer you back to your analysis software vendor for help with their applications. ✦

Faster Network Access Now Available

We've been working with Tymnet (MCI) to upgrade network access for our U.S. and Canadian customers. We are pleased to announce that we have increased the top speed for Tymnet access from 2400 to 9600 baud. Network users who wish to take advantage of the faster speed should call MCI's service line at 1-800-628-3497. They have a touch-tone menuing system, from which you should select "Network Access Inquiries."

When the representative comes on the line, simply request the local 9600 baud access number from your location. The new number should be inserted on page 1 of User Constants, where the Baud Rate should also be changed to 9600.

We are still negotiating with the networks for lower rates, so the promised reduction in network premiums is not yet in effect. Watch this journal for more information on changes to our month-by-month rate schedule. ✦

Holiday Schedule

CSI will be closed for voice communication on Monday, September 4th for the Labor Day holiday. Data from those exchanges that remain open will be available as usual and the CSI host computer will be accessible 24 hours per day throughout the holiday weekend.

CSI Software Product Summary

Please check all that apply and complete the information box at right.
Mail or fax to CSI, 200 West Palmetto Park Road, Boca Raton, Florida 33432; Fax: (407) 392-7761

❑ **QuickTrieve®/QuickManager®** for PC - To retrieve, manage & edit data (includes 1995 Alerts Calendar); New daily user $59. QuickTrieve/QuickManager version 4.06 upgrade (for current QuickTrieve users only): $39; shareware demo disk $5

❑ **1995 Commodity Alerts Calendar** for use with QuickTrieve $20; Calendar upgrade for current QT 4.06 users $10

❑ **QuickPlot®/QuickStudy®** for PC - Charting & analysis software (requires QT/QM) $89

❑ **Trade Data Manager™** - Macintosh downloader & accounting program $59; upgrade $49 or *FREE* with $100 history order

❑ **Trading System Performance Evaluator™** (TSPE) for PC - Computes your system's capital requirements $149

❑ **Trader's Money Manager™** for PC - $399 (includes TSPE); Demo disk: $15

❑ **TraDe$k™** for PC - Traders' complete accounting system - CSI daily user $149; Unrestricted use $299; 30-day trial version $22

❑ **Seasonal Index Value Pack** for PC - Ten years of history for 33 popular commodities $315

❑ **Daily Updates** for PC - Starting at $10.80 per month

❑ **CSI Technical Journal Subscription** - $35/Yr. - Reprints $5/each Issues requested: _____

❑ **CSI Mailing List** - $200/1,000 names (CSI users omitted)

❑ **CSI Product Catalog** -*FREE*

Please add $29 per software package for overseas shipping.

NAME _____

ADDRESS _____

_____

_____

DAY PHONE (____) _____

USER ID# _____

DISK PREFERENCE
☐ 5.25"/360K ☐ 5.25"/1.2 MB (HIGH DENSITY)
☐ 3.5"/720K ☐ 3.5"/1.44 MB (HIGH DENSITY)

METHOD OF PAYMENT (PREPAYMENT REQUIRED)
☐ CHECK ☐ MASTERCARD ☐ VISA
☐ DISCOVER ☐ AMERICAN EXPRESS

AMOUNT ENCLOSED $ _____

CARD # _____

EXP. DATE _____

SIGNATURE _____

9/95

All prices subject to change without notice.

The Missing Analysis...

neural network, so you have to be careful on this one. If you know a given pair of inputs are correlated, then the ratio of the correlated series may be substituted. On the other hand, it may be possible to represent the series differently so the model will not focus on this aspect. Again, the preprocessing step needs to address such problems so the resulting model will not be searching for a unique solution when perhaps several less than viable (unwanted) solutions exist.

Ignore the preprocessing step and your trading model could be a nightmare that will train well but will not deliver post-training profits. Return to the first couple of paragraphs of this discussion and read them again. The novice avoids the preprocessing step. If you attempt to omit it, the modeling step will take months longer and it probably won't work when you believe you have finally got it. If you don't get it, don't try it or the market will take it. (The last "it" was your money.)

Next month we'll discuss building a model that will help you turn your next great idea into a system that will consistently return profits from trading the market. No matter where you go from here in your journey, remember that the best starting point for all market analysis is clean data. Remember also that raw data is not the most important cog in the neural network machine. Preprocessed data is the ultimate food for a trading system that performs well in hindsight and in real time. ✦

Bob Pelletier

Market Statistics Update

NAME, SYMBOL & EXCHANGE CHANGES

2413	ALMO	Alamo Group Inc - New exchange: NYSE; New symbol: ALG.
12078	APPS	Apps Dental Inc - New symbol: CPDN.
1807	ASTI	Astrum Int'l Cp - New name: Samsonite Cp; New symbol: SAMC.
4499	BZHKF	Belize Holdings Inc - New name: BHI Cp; New symbol: BHIKF.
8214	BIND	Bindley Western Industries Inc - New exchange: NYSE; New symbol: BDY.
1517	CLEBC	Celebrity Entertainment Inc- New symbol: CLECD.
8419	CFCN	Commercial Federal Cp - New exchange: NYSE; New symbol: CFB.
4101	CPTD	Computer Data Systems Inc - New symbol: CDSI.
12167	CRON	Cooper Cameron Cp - New exchange: NYSE; New symbol: RON.
6395	FBF	CS First Boston Income Fund Inc - New name: BEA Income Fund Inc.
6318	FBI	CS First Boston Strategic Income Fund Inc - New name: BEA Strategic Income Fund Inc.
4548	DVS	Davstar Industries Ltd CL A - New name: Urohealth Systems Inc Class A; New symbol: URO.
8584	EZEMA	E-Z-EM Inc CL A - New symbol: EZMA.
7187	FCA	Fabri-Centers of America Inc - New name: Fabri-Centers of America Inc CL A; New symbol: FCAA.
1388	LION	Fidelity Southern Cp - New name: Fidelity National Cp.
7803	FCX	Freeport-McMoran Copper & Gold Inc - New name: Freeport-McMoran Copper & Gold Inc. CL A; New symbol: FCXA.
3250	GUN	Gundle Environmental Systems Inc - New name: Gundle SLT Environmental Inc.
1129	HGI	Horizon Outlet Centers Inc - New name: HGI Realty Inc.
3290	HTK	Howtek Inc - New exchange: OTC; New symbol: HOWT.
2086	INDE	Independent Telemedia Group Inc - New name: Indenet Inc.
2948	IMTAD	Information Management Technologies Cp CL A - New symbol: IMTKA.
7312	IS	Interstate/Johnson Lane Inc - New symbol: IJL.
9005	IWBK	Interwest Savings Bank - New name: Interwest Bancorp.
1753	LFIIF	Laser Friendly Inc - New symbol: GLCCF.
2815	MADGF	Madge, N.V. - New name: Madge Networks N.V.
5946	NACC	National Auto Credit Inc - New exchange: NYSE; New symbol: NAK.
4413	NETX	Natutal Earch Technologies Inc - New symbol: USHG.
1916	APHC	Partech Holdings Cp - New name: Tropic Communications; New symbol: RPC
7937	PRAI	Peer Review Analysis Inc - New name: Core Inc; New symbol: CORE.
12002	PSTB	Perpetual Savings Bank - New name: Perpetual State Bank.
3446	JMP	Peters (J.M.) Co - New symbol: CPH.
9497	QLTIF	Quadra Logic Technologies Inc - New name: QLT Phototherapeutics Inc.
6882	RRRR	Renaissance Communications Cp - New exchange: NYSE; New symbol: RRR.
3732	DRZ	Sterling Healthcare Group Inc - New symbol: STER.
2576	VVTVA	Valuevision Int'l Inc CL A - New symbol: VVTV.
7975	VITA	Vitamin Specialties Cp - New name: Healthrite Inc; New symbol: HLRT.
9949	WILLA	Wiley (John) & Sons Inc CL A - New exchange: NYSE; New symbol: JWA.