CSI TECHNICAL JOURNAL

DATA ✦ SOFTWARE ✦ SERVICE

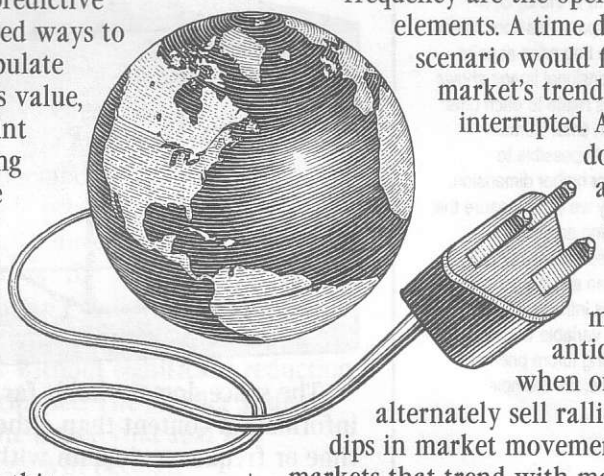
Volume XI, Number 10 *October 1995* *$10.00*

In This Issue

Editor: Sabrina Carle
Publisher: Commodity Systems, Inc.

The Care and Feeding of a Market Trading Model

In the last issue of the CSI Technical Journal, we addressed the need for data preprocessing in model-building exercises. We stated that the most important requirement of any model building task is to provide flawless data preprocessed into a form that will enhance and foster predictive potential. We suggested ways to transform and manipulate input to maximize its value, and we made the point that ignoring or giving light attention to the data preprocessing step will curb your model's effectiveness when anticipating market movement.



Without losing sight of the importance of the data preprocessing step, in this issue we will move forward with the next analytical requirement by exploring a few model building approaches using preprocessed time series data. Mark Jurik, in his book Virtual Trading, discusses categorizing time series with reference to the space domain, the time domain and the frequency domain. Jurik contends that when a time series can be expressed with respect to its relevance to other time series at points in time, it lies within the space domain. When a time series is expressed with reference to itself over time in a trending context, it is in the time domain. Finally, when a time series is expressed with reference to itself in a repetitive or periodic frequency context, it is in the frequency domain. We like his approach and will use these ideas as a springboard for further discussion on model building.

Time and Frequency Domain Models

It is common for model builders to use approaches where time and frequency are the operative elements. A time domain scenario would follow the market's trend until it is interrupted. A frequency domain approach to systems development might anticipate when one should alternately sell rallies and buy dips in market movement. Mixing markets that trend with markets that are choppy is an exercise that could involve both the time domain and the frequency domain simultaneously.

Applying trend following techniques to one class of markets and anticipating peaks and troughs on another class are two dissimilar approaches which could be applied together. Such an application would be a good way to jointly capture the effects of both forms of market movement. This combined approach lets you benefit from a diversity in market type and a diversity of analytical method. If the equity curves of pairs of differing classes (market type and analytical method) are uncorrelated and jointly profitable, risk and investment costs may be minimized

(continued on Page 2)

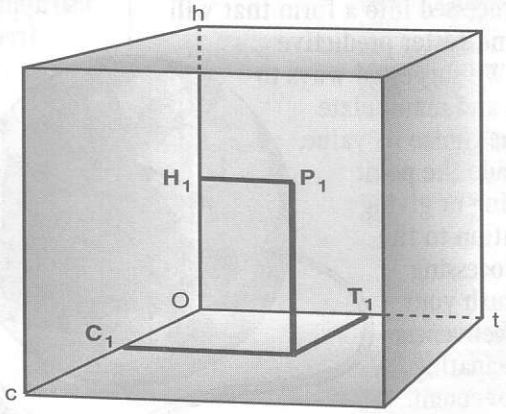
The Care and Feeding...

(continued from page 1)

while profits are maximized.

Space Domain Models

Even though analysts tend to favor approaches in the time and frequency domains, these may not be the most lucrative or the most reliable forms of mathematical system development. Much more information can be derived from the space domain. In the space domain, related markets (transformed into relative terms and introduced as ratios or differences) can be measured to determine their influence.



The space domain illustration to the right shows how time (O to t) in the c,O,t plane relates to the price of corn (O to c) and hogs (O to h) in the C, O, h plane. Point P_1 in space ties corn's price at point C_1 with the hog price at point H_1. Our three dimensional image shows how corn and hogs relate to each other through changes in time. Human limitations make it impossible to visualize a fourth or higher dimension, but mathematically we can measure the effects of introducing additional dimensions of independent input. In other words, we can still measure P_1 as more information is introduced. P_1 may be a performance variable we hope to process in predicting future price of some other variable, for example.

The space domain holds far more information content than either the time or frequency domain without necessarily consuming excessive freedom restricting control. In the space domain, elements (the effects of individual time series on performance) in multidimensional space at each point in time can be weighted. The control can then be spread in a linear or curvilinear fashion through coefficients that range from zero to one. The weighting would be applied to various time series elements as required. All independent variable groupings would be weighted and the sum of the weights should be normalized to add to one, for example. An iterative process could be employed in which weights would be computed with a goal to maximize the profit for the series you wish to predict.

The above model would also accumulate profits and losses in dollar units of the dependent variable based upon relative movement of the two or more independent variables. The independent variables here would be raw series or ratios and differences of raw series. The resulting profits and losses should be placed in a relative form such as by using a Sharpe ratio and a percentage return on investment or equivalent. This way different markets can be compared to assess and rank relative merit.

The Dependent Variable

Feeding your model remains a considerable task which involves gathering relevant information as input. Your first step, of course, is identifying the time series element you wish to predict. Since we are talking about a market tool for prediction, your time series must be something which is tradeable and which has substantial historical information on a daily basis. The longevity of the data base information you will need to derive a responsive model will increase as you introduce more parametric control.

Limiting Parametric Control

The size of the data base in terms of the number of records you will process is important and should be balanced with the number of parameters you plan to introduce. Actually, it is the number of resulting trades (samples) that are traded off with parameters, but because the number of trades would go up as the period is extended, increasing the size of your data base has an equivalent effect.

Suppose you wish to fit a curve through a scattering of points over time in anticipation of predicting one or two days into the future. You could use the mathematical system of least squares and consume one parameter by fitting a straight line through the data and extrapolate ahead. If you want to go further by adding some curvature,

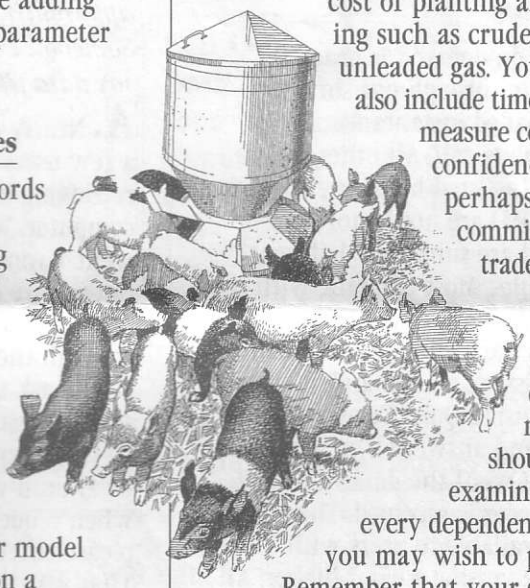
you might try a second degree polynomial, etc. For every additional degree you introduce, you would be adding another restrictive control parameter to the process.

Balancing Parameter Counts With Input Samples

The number of daily records you introduce will serve to temper the effects of adding parameter control, but only up to a point. You will find that by minimizing parameter control, your model will show more and more promise when passed against test data. It is fairly simple to find parameter settings that will make your model perform well in hindsight on a training period. Getting your model to perform in the future has to do with both model design and the method by which you introduce and pre-process your inputs. The scope and variability of your input is also an important factor.

Searching For Relevant Model Input

Once you have identified the time series to be predicted (the dependent variable), the space domain requires that you identify relevant inputs which will influence future movement. Last month's article discussed broad categories to be incorporated into your pre-processed data. We would like to expand on these with more specific examples. If your dependent variable is corn, for example, consider adding independent variables in related commodities that grow on the same land such as soybeans and wheat. Also consider derivative commodities of corn (products developed from corn) such as live hogs or live cattle and substitute commodities like oats and soybean meal. For a broader look at the financial picture, add series that affect the cost of money such as T-Bonds, T-Bills and the Federal Funds rate. Don't overlook commodities that affect the cost of planting and harvesting such as crude oil and unleaded gas. You might also include time series that measure consumer confidence, and perhaps the CFTC's commitment of trader's statistics.



A set of influential, independent markets should be examined for every dependent variable you may wish to postulate.

Remember that your goal is to get reliable predictions from the data set you introduce. Once you have a good start on the prediction problem, try to simplify the process by searching for marginal inputs that can be removed without significant reduction in performance. The simpler your model, the better your real life predictions will become.

Input Longevity

Neural Network researchers often state that six months to two years of daily data is sufficient for a typical study, but such a limited data base may not hold sufficient lessons of the past to produce reliable future predictions. Computer speed is usually an important restriction on input longevity and scope. Most analysts would gladly analyze two decades of daily data if their computer could make sense of the input in a timely manner. Unfortunately, a staggering amount of compute time is required for applying the perhaps ten or more parameters that may be relevant to drive a given model. The number of combinations of events to search becomes formidable

(continued on Page 6)

"Getting your model to perform in the future has to do with both model design and the method by which you introduce and pre-process your inputs."

Ask Customer Service

Each month on this page our Customer Service staff addresses common questions on topics of interest to many CSI subscribers. This month our history on-demand data delivery mode takes the spotlight.

Q. *What is the difference between history on-demand and other types of historical data?*

A. History on-demand is ordered electronically through our software and downloaded instantaneously in the same phone call. All other means of delivery (assisted by phone, on disk or on CD-ROM) are not history on-demand and are subject to different price schedules. Most CSI data, with the exception of options and seasonal index series are available on-demand.

To get history on-demand, PC users must order using QuickTrieve's® Order Subsystem and answer (Y)es when asked if you want the data transmitted when the order is received. This service is available to users with QuickTrieve version 4.05 or higher. All phone histories transmitted through our Macintosh software are now processed as history on-demand.

Q. *I found some of the files I ordered as history on-demand were incomplete. Although I requested a long history, some files contained only one day of data. How can I correct this problem?*

A. Missing data in on-demand history is sometimes caused by modem incompatibilities. A simple change to the modem initialization string may correct the problem. From the QuickTrieve main menu, select H) CHANGE USER CONSTANTS. Then press <Pg Dn> until you reach Page 7. The fist line there is Modem Init. Before making any changes, please print the screen or write down your current entry, then change it to the following:

Modem init: AT&D2V1Q0N0S37=0 (All three ovals in the above line are zeros, not the letter O.)

Press <Esc> to exit and be sure to save your changes as prompted. This should eliminate the mysterious missing data from on-demand history orders.

Q. *I spent a lot of time typing a large history on-demand order which was apparently accepted by your computer. Somehow I ended up with nothing in my data files. What caused this?*

A. Nearly every day we get calls from a few users who were unsuccessful at retrieving history from our host computer. We try to diagnose what went wrong on a case-by-case basis, but we are often hampered by a lack of information on what was requested through the software.

We ask that you take the following steps to assure that your requests will be filled promptly:

1) Until you get used to the system or when requesting a new type of data, please enter only a few lines per order. That way, if an error prevents us from supplying the data, you won't have wasted much time or effort.

2) Review your order before transmission to make sure that all the entries are correct. Common errors include: stocks identified as commodities (or vice versa); both CSI number and symbol supplied when either number or symbol is required; delivery month and delivery year reversed; absence of valid start dates for stocks, Perpetual Contract,® cash or other continuous series; start and end dates reversed; 0 for delivery month for cash or Perpetual Contract data; etc. Any of these entries would result in a "file assign error" and no data would be supplied.

3) Before transmitting your order, please use the PRINT WORK DONE THIS SESSION option from the order subsystem menu to make a printout. If there should be a problem, you'll have it to review and, if necessary, re-enter. Our service representatives will ask you to read or fax your entries to help determine what went wrong. This is an important safeguard because once your order is transmitted correctly, it

(continued on Page 6)

Faster Network Access a Hit!

Since announcing last month that we now offer 9600 baud access via Tymnet (MCI), many subscribers have switched to the faster service. 9600 baud retrieval seems to be coming off without a hitch. We urge everyone using network access to switch to the 9600 baud Tymnet alternative.

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 "2-Network Access Inquiries." When the representative comes on the line, simply request the local 9600 baud access number for your location. The new Tymnet phone number should be inserted on page 1 of User Constants, where the Baud Rate should also be changed to 9600.

Lower Tymnet charges have enabled us to reduce our the network surcharge by 15% or more for custom portfolio users paying on a month-by-month basis. The lower costs are effective immediately, but will not be reflected in your current invoice. November invoices will show a prior month service adjustment for October, making the lower prices retroactive to October. ✦

New Customer Service Hours

Our customer Service hours are now 8 a.m. - 10 p.m. weekdays and 9 a.m. - 1 p.m. Saturdays. The Customer Service telephone number is (407) 392-8663. Customer Service support is not available through our toll-free marketing line.

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
- ❑ **TraDesk™** 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 _____

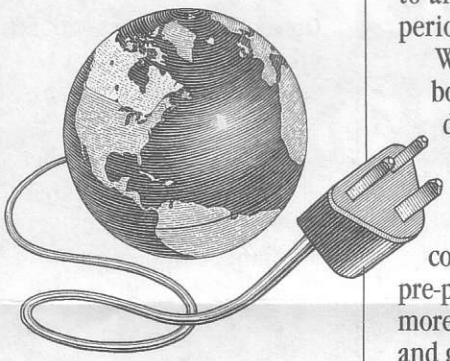
SIGNATURE _____

10/95

All prices subject to change without notice.

The Care and Feeding

(continued from page 3)



as the training period is extended. The result is often a compromise that leads to an abbreviation of the training period.

We heartily recommend Mr. Jurik's book Virtual Trading, which probes deeply into many analytical ideas that are important to any model building exercise. When sound model-building techniques are combined with properly selected and pre-processed input data, the result is more reliable trade recommendations and greater profits. ✦

Bob Pelletier

Reference: Virtual Trading Mark G. Jurik Probus Publishing Co, 1995, Chicago

Ask Customer Service

(continued from page 4)

is removed from the screen.

4) Please call us if you have a problem with any sort of data retrieval. The sooner we can pinpoint a problem - at your end or ours - the sooner we can solve it.

Q. *I ordered some cash data through the order subsystem and found the same value in all four price fields for every trading day. How can I get the full daily range and the daily price fluctuations?*

A. This is a glitch we thought we had corrected earlier. It happens only occasionally on a seemingly random basis. As yet, we aren't sure why. Anyone receiving code 54 cash data as history on-demand that shows up as just a closing value in all four price fields with the same prices every day should call customer service for a free replacement. We hope to correct this problem by the time you receive this journal. We regret the inconvenience. ✦

Market Statistics Update

NAME, SYMBOL & EXCHANGE CHANGES

8044	ATCC	Aintran Cp - New name: Mesaba Holding Inc.
1899	ANMC	Analytical Nursing Management Cp New Symbol: AMEX
12290	CNLDV	Conolog Cp WI - New symbol: CNLGD.
12291	CNLWV	Conolog Cp CL A WT - New symbol: CNLGW.
2892	CNUUC	Continum Group Inc - New symbol: CNUDC.
7295	INT	Int'l Recovery Cp - New name: World Fuel Services Cp.
7312	IS	Interstate/Johnson Lane Inc - New symbol: IJL.
4027	KNIC	Knickerbocker (L.L.) Co, The - New symbol: KNIDV.
12566	MPML	Montana Precision Mining Ltd - New name: MPM Technologies Inc.
2858	MLTI	Multicare Cos - New symbol: MUL; New exchange: NYSE.
12587	NRGIA	National Energy Group Inc CL A - New symbol: NEGX.
9289	NELL	Nellcor Inc - New name: Nellcor Puritan Bennett Inc.
3030	LUXY	Nickelodeon Theater Co - New name Cinemastar Luxury Theaters.
12596	OAKC	Oakhurst Capital Inc - New name: Oakhurst Co.
4352	PARQ	Parcplace Systems Inc - New name: Parcplace-Digitalk Inc.
1916	APHC	Partech Holdings Cp - New name: Tropic Communication
7470	PPC	Patrick Petroleum Co - New name: Goodrich Petroleum Cp.
8081	SLVRF	Silverado Minos Ltd - New symbol: GOLDF.
4481	SOLQD	Solo Serve Cp - New symbol: SOLOQ.
2832	UPIX	Unapix Entertainment Inc - New exchange: AMEX; New symbol: UPX.
6875	VFIC	Verifone Inc - New exchange: NYSE; New symbol: VFI.

STOCK SPLITS

2830	ATRM	Aetrium Inc	3/2	950828
4025	ABII	Amer Business Information	3/2	950815
5984	AMGN	Amgen Co	2/1	950816
6868	ATML	Amtel Cp	2/1	950809
2609	AINN	Applied Innovation Inc	2/1	950901
7033	BEZ	Baldor Electric Co	3/2	950907
8161	BMCS	BMC Software Inc	2/1	950815
3113	CAV	Cavalier Homes Inc	5/4	950816
3725	CTN	Centennial Technologies	3/2	950831
8409	COMR	Comair Holdings	3/2	950811
8413	CMDLD	Comdial Cp	1/3	950808
5566	CA	Computer Assoc Intl Inc	3/2	950906
2892	CNUDC	Continum Group Inc	1/4	950814
2068	CREE	Cree Research Inc	2/1	950816
2514	DAVD	Davidson & Associates Inc	2/1	950907
8603	ELRC	Electro Rent Cp	3/2	950822
7187	FCAA	Fabri-Centers of Amer Cl A	2/1	950817
8680	SRCE	First Sources Cp	3/2	950822
1610	FKBC	First Knox Banc Cp	2/1	950905
2806	IDBEF	ID Biomedical Cp	2/1	950808
1538	INSO	Inso Cp	2/1	950905
9607	ISSS	Intergrated Silicon System	2/1	950825
2434	INTU	Intuit Inc	2/1	950822
4832	IROQ	Iroquois Bancorp Inc	2/1	950901
2719	JKCLD	Jockey Club	1/20	950830
4027	KNIDV	Knickerbocker Co	5/1	950831
9111	LLTC	Linear Technology Cp	2/1	950905
9068	LYTS	LSI Lighting Systems Inc	3/2	950816
3662	MSL	Midsouth Bancorp Inc	4/3	950905
5876	MYL	Mylan Labs Inc	3/2	950816
4700	NKOT	Nu-Kote Holdings Inc	2/1	950816
8440	ICTL	Ocom Cp	4/3	950814
7469	PKE	Park Electrochemical	2/1	950816