SPlus and S+FinMetrics

An advanced Toolkit for
Time Series Modeling and
Quantitative Analytics
S+ FinMetrics

Command-line based SPlus module
500+ functions for
  + financial time series analysis,
  + econometric modeling and prediction
  + advanced visualization capabilities

Created with much support of academics for
  + researchers and practitioners in the finance industry
  + academic researchers graduate students in economics and finance
• Eric Zivot, Jiahui Wang
• Alexander McNeil (EVIS & library QRMLib)
• Rene Carmona & Julia Morrison (EVANESCE)
• Siem Jan Koopman (SsfPack C library)
• Rollant Gallant & George Tauchen (SNP and EMM Fortran lib)
• Jan Beran et al (long memory functions)
  and many more ...

• Cooperation with academia essential for Insightful:
  + SPlus (Insightful) 150 employees
  + Matlab (Mathworks) > 1500 employees
  + SAS > 10000 employees
Key features of SPlus
Key features of SPlus

- SPlus Server
- SPlus workbench
- I-MINER
- FinMetrics
- NuOpt
- many libraries: wavelets, robust stats, ...
- FAME S+connector
- ability to import data from
  ODBC connection, SQL, ORACLE, ACCESS, EXCEL, SAS, SPSS, MATLAB, GAUSS, LOTUS, MINTAB, SYSTAT,...
SPlus Server

• Server-based version of SPlus
• Deployment of applications developed in SPlus using a client/server architecture
• Adds features to desktop SPlus for Web integration, JSP and ASP APIs, and batch processing
• Supported on multi-CPU Unix and Windows Server architectures
• Allows unlimited use of the SPlus language engine and analytics on licensed CPUs
Insightful MINER

- handle large datasets
- drag-and-drop functionality
- easy creation of visual workflows for sophisticated models
- ability to run SPlus scripts using all modules

- self documenting
- easy to deploy, adapt, pass on
Step 1: Connect to Data Base

1. Load Term Structure Data

Step 2: Descriptive Analysis

2. Correlation Matrix
3. Descriptive Statistics
4. PCA Factor Model
5. 2D Term Structure
6. QQ Plot

Step 3: PCA Modelling

Step 4: Select Forecasting Model

Step 5: Backtest Forecasting Model

Step 6: Factor Modelling

Step 7: Monte Carlo Simulation

Step 8: Stress Testing

Step 9: Pricing Step
SPlus Workbench

- plugin for open source Eclipse
- syntax highlighting
- menu selections for source control
- debugger, ability to set breakpoints
- monitor memory usage
- integrated command line console
- embedded graphs
NuOpt (SPlus module)

Numerical optimization software package

- linear programming problems
- mixed integer programming problems
- transportation problems
- quadratic programming problems
- unconstrained nonlinear optimization problems
- constrained nonlinear optimization problems
- multi-objective programming problems
Portfolio optimization with NuOpt

- optimal portfolios beyond mean variance approach
- basket selection with turnover constraints
- cardinality constraints (number of names constraints)
- round Lots: transactions are allowed in integer multiples of round lots
- Lower partial moments optimization
- Scenario optimization

➢ NuOpt in combination with S+Bayes and robust library
  - robust statistical methods for portfolio construction
  - Black-Litterman models
FAME S+Connector

• Partnership with Sungard
• Combine Sungard’s FAME data management solution for storing and managing real-time and high-volume time series data to the power of SPlus

➤ provide access to FAME’s comprehensive data sources
➤ exposes the power of FAME’s in-database manipulations to SPlus
➤ provide cross platform support
Benefits of S+FinMetrics?

+ Consistent syntax and ease of use
  - All modeling functions make use of formulas, which makes mode specification an easy task; least input is required from the user; automatic model selection for many time series models, such as ARIMA, FARIMA, SEMIFAR, VAR, etc.
+ Full support for using timeSeries objects
  - Can use timeSeries data frames for modeling functions involving formulas; time series leading and lagging functions for dynamic regression; missing value interpolation; etc.
+ Commonly requested econometric features
  - White correction or Newey-West correction for time series regression; seemingly unrelated regression; unit root testing; co-integration and vector error correction models; state space modeling; etc.
+ Facilitate simulation and forecasting
  - All time series models can be simulated and forecasted; using simulations to obtain confidence intervals for forecasts.
+ Direct Financial Applications
  - Factor modeling, yield curve fitting, backtesting models,
  - asset return forecasting, volatility modeling, VaR, risk modeling
Functionality

- Time Series Manipulation and Visualization
- Rolling Statistics and Technical Indicators
- Rolling Estimation and Backtesting Strategies
- Time Series regression (static, dynamic regression)
- Systems of equations estimation (SUR)
- Classical Time Series Analysis
  (ARIMA, ARIMA, spectral analysis, unit roots ...)
- Long-memory time series
- Vector Autoregressive models (VAR)
- Cointegration and Vector Error Correction Model (VECM)
• Univariate and multivariate GARCH Volatility Modelling
• State Space Modelling
• Factor Models for Asset Returns
• Fixed Income Analytics (rate conversion, yield curve fitting ...)
• Extreme Value Theory and Copula Analysis
• Robust change detection (robust REGARIMA ...)
• GMM
• Nonlinear Times Series Models (Markov switching and threshold models)
• Continuous Time Models (approx. & simul. SDE)
• Semi nonparametric Conditional density models
• EMM
Sources of information

• On the web:
  + Insightful Finance
    http://www.insightful.com/industry/finance/default.asp
  + FinMetrics product description
  + FinMetrics datasheet
  + FinMetrics feature list
  + Documentation

• With the software:
  - function guide
Forecasting: Insightful's comprehensive suite of statistical functions provide the tools for building the right model faster, or adjusting your existing models to forecast new market realities.

Derivative Pricing: Access a complete suite of computation tools, plus support for a wide range of random distributions used in Monte Carlo pricing, ideal for pricing options and derivatives.

Back-testing: Easily test portfolio construction or trading rules on historical data and statistically analyze performance.

Value at Risk: Utilize a broad range of statistical tools including GARCH tools and Extreme Value Theory for estimating volatility and forecasting VaR.

Program Trading: Employ our professional services team to create a trading platform that combines real-time quotes from multiple markets with the analytics to determine buys, sells and amounts that meet your toughest specifications.
Risk Management Applications

Credit Risk Modelling: Implement advanced industry-leading analytics and model portfolio losses using the Merton-based approach.

Capital Management: Use both an aggregated and a drill-down view of economic capital to make more informed decisions.

Risk Aggregation: Easily access an aggregated view of risk for your organization by line-of-business, geography, and risk type.

Scenario Analysis: Analyze buy/sell decisions with what if scenarios which can be quickly assessed and result in rapid decision making.
What’s New in S+FinMetrics 3.0

(S+FinMetrics 3.0 will most likely be released this year)

+ Extended State-Space Functionality
+ Extended GARCH Functionality
+ Extended Monte Carlo and Copula Functionality
+ High-Frequency Functionality
+ Option pricing
+ Fixed-income calculations
+ Cash flow functions
+ Enhanced big data functional support
Simple Descriptive Tools

- Time series manipulation/smoothing/filtering/aggregation
- Technical Analysis
- ARIMA with Regressors and Long Memory
- Dynamic Time Series Regression
- Tests for Unit Roots, Cointegration, Nonlinearity
- Extreme Value Distributions
- Simulate Solutions to SDEs
- Cash flow functions
- Nonlinear regime switching and neural networks
- General Rolling Estimation
- Seemingly Unrelated Regression
- High-Frequency Functionality
- Vector Autoregression and Cointegration
- GARCH – Univariate and Multivariate
- Method of Moments Estimation – GMM & EMM
- Copulas
- State Space Models and Kalman Filter Tools
- Statistical Factor Models for Large Portfolios
- Yield curve modeling
- Fixed-income calculations
- Option pricing

New or enhanced feature in FinMetrics 3.0

Advanced Modeling Tools
Time Series Manipulation and Visualization
SPlus “timeSeries” Objects

+ Combines data with “timeDate” object
+ Flexible enough to describe essentially all types of financial time series data
  - Regularly spaced calendar data
  - Irregularly spaced tick-by-tick data
  - Allows time-zone specification
  - Easy event handling: Holidays, market closures, etc.

+ Powerful plotting functionality
  - Manipulation
    - Lags
    - Returns
    - Aggregation / Disaggregation
  - Visualisation
    - Trellis QQ Plots
    - Trellis Histograms
    - Trellis Residual Plots
Time Series Plots

Monthly prices on Microsoft and S&P 500 Index

Returns data

Copyright ©1999-2007
Insightful Corporation. All Rights Reserved.
Multivariate Time Series

Prices on Dow Jones 30 assets

Dow Jones 30

Dow Jones Prices

DIS
WMT
UTX
SBC
PG
MO
MMM
MSFT
MRK
MCD
JNJ
JPM
IP
IBM
INTC
HON
HD
HWP
GM
GE
XOM
EK
DD
KO
C
CAT
BA
T
AXP
AA

Copyright ©1999-2007
Insightful Corporation. All Rights Reserved.
Motivation

+ assess the model’s stability over time.
+ “poor man’s” time varying parameter model.
+ “backtest” a statistical model on historical data to evaluate stability and predictive accuracy.
Technical Indicators
• Describe the activity of financial instruments and obtain trading signals
• Mostly based on moving average techniques
• Broadly, the main technical indicators can be classified into four categories:
  + price indicators
  + momentum indicators and oscillators
  + volatility indicators
  + volume indicators
Momentum Indicators

Moving Average Convergence Divergence

Differences from signal

Buy Signal
Sell Signal
Volume Indicators

Typical Price of Microsoft Data

Stock Volume

Chaikin's Volatility Indicator

Accumulation / Distribution Index
Regression Modeling
Predicting Stock Returns

![Graph showing stock index values over time, with 'msft' indicating Microsoft Corporation's stock. The x-axis represents the index values, ranging from 1480 to 1580, while the y-axis shows the values from -0.05 to 0.10. The graph highlights fluctuations in stock returns over the specified period.]
Rolling Estimation and Back Testing
Motivation

+ assess the stability of the model’s parameters
+ provide a simple “poor man’s” time varying parameter model
+ backtest a model’s predictive performance with rolling predictions
+ compare the predictive accuracy of rival models
Prediction and Backtesting

Returns on MSFT and 1-step forecasts

1-step forecast error

<table>
<thead>
<tr>
<th>Month</th>
<th>Actual</th>
<th>1-step Forecast</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mar</td>
<td>-0.04</td>
<td></td>
</tr>
<tr>
<td>Apr</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>May</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>Jun</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>
Backtesting Risk Measures

Value at Risk Backtesting

90.8% points within VaR Bound
System of Equations
Residuals versus Time

Seemingly Unrelated Regressions
Classical Time Series Analysis
Multivariate Series: `ret.ts`

- **MSFT**
- **MSFT and SP500**
- **SP500 and MSFT**
- **SP500**

**Autocorrelation Plots**
Long Memory Modeling

Long Memory Modeling
Series : abs(sp500)
Vector Autoregressive Models
VAR Plot Method
VAR Prediction

Classical VAR Prediction

Values

index
Bayesian VAR Prediction
Cointegration Analysis
Simulated trivariate cointegrated system

Cointegrated residuals
Volatility Modeling (GARCH)
GARCH models

Classical time series models focus on conditional first moments
  + Autoregression
  + Moving average

These models do not take into account stylized facts about finance data:
  + Volatility clustering
  + Fat tails
  + Volatility mean reversion

GARCH models do.
Volatility Clustering in Financial Time Series
• **Univariate GARCH**
  + ARCH, GARCH
  + EGARCH, TGARCH, PGARCH, GARCH-M
  + exog. variables in conditional mean or variance eq.
  + non-Gaussian error (t-distr., GED)
  + fitting, prediction, simulation

• **Multivariate GARCH**
  + EWMA, DVEC
  + matrix diagonal, BEKK, CCC, PC
  + exogenous variables and arma terms
  + conditionally t-distr. residuals
  + diagnostics (JB, SW, LB, LM)
  + prediction & simulation Modelling extreme values
GARCH Models

Series and Conditional SD

Original Series

Conditional SD

Values

-0.1 0.0 0.1

0.02 0.03 0.04 0.05

Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1

State Space Modeling
Based on a library SsfPack
  + Kalman filter/smooth, forecasting, simulation

Applications:
  + Estimating time-varying parameter regression
  + Affine term structures models

General time series modelling framework
Useful in model simulation
Kalman Filters

Disturbance Simulation

Draws

state

response

Copyright ©1999-2007
Insightful Corporation. All Rights Reserved.
Factor Models for Asset Returns
• Statistical factor models for returns & principal components
• Macroeconomic- and fundamental factor models
  (regression models)
• Explaining the variation in a reduced number of dimensions
  (dimension reduction techniques)
• Produce stable estimates for covariance structure
• Can be used in combination with S+NuOpt for optimal asset allocation and portfolio trading strategies
Factor Models

Top Positions in Factor 1 (Percentages)

<table>
<thead>
<tr>
<th>Company</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBY</td>
<td>0.224</td>
</tr>
<tr>
<td>SCON</td>
<td>0.256</td>
</tr>
<tr>
<td>PUMA</td>
<td>0.275</td>
</tr>
<tr>
<td>MT</td>
<td>0.294</td>
</tr>
<tr>
<td>FNV</td>
<td>0.309</td>
</tr>
<tr>
<td>OWC</td>
<td>0.320</td>
</tr>
<tr>
<td>MT</td>
<td>0.340</td>
</tr>
<tr>
<td>FNV</td>
<td>0.357</td>
</tr>
<tr>
<td>OWC</td>
<td>0.378</td>
</tr>
<tr>
<td>MT</td>
<td>0.397</td>
</tr>
<tr>
<td>FNV</td>
<td>0.406</td>
</tr>
</tbody>
</table>

Variances:
- F.1: 0.167
- F.2: 0.224
- F.3: 0.256
- F.4: 0.275
- F.5: 0.294
- F.6: 0.309
- F.7: 0.320
- F.8: 0.340
- F.9: 0.357
- F.10: 0.378
- F.11: 0.397
- F.12: 0.406

Copyright ©1999-2007
Insightful Corporation. All Rights Reserved.
Term structure of Interest Rates
• Rate conversions

(convert between discount-, forward- and spot rates)

• Fit term structure of interest rates using
  + Quadratic- and cubic spline regression
  + smoothing spline fitting
  + Nelson-Siegel- or Nelson-Siegel-Svensson function
U.S. Yield Curve with Svensson Function: April 1987
Extreme Value Theory and Copulas
• **Tail estimation for more robust quantiles**
  + Generalised Extreme Value Distribution
  + Generalised Pareto Distribution
  + Tail estimation by POT

• **Copulas: estimation and simulation of multivariate distributions**
  + Parametric copula classes and families
  + Estimation
  + Simulation
  ➢ Compute Risk Measures: VaR, ES
Tailplot for Pareto Fit

- $x$ (on log scale)
- $1-F(x)$ (on log scale)

99 95
99 95

Extreme Value Analysis
Fitting Copula on 2-D fat-tailed data

copula fit vs. observed

simulated
actual
Derivative Pricing
Calculation of Vega for various scenarios

<table>
<thead>
<tr>
<th>Volatility: 50%</th>
<th>Volatility: 100%</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="3D Graph" /></td>
<td><img src="image2" alt="3D Graph" /></td>
</tr>
<tr>
<td>Volatility: 10%</td>
<td>Volatility: 25%</td>
</tr>
<tr>
<td><img src="image3" alt="3D Graph" /></td>
<td><img src="image4" alt="3D Graph" /></td>
</tr>
</tbody>
</table>
Ornstein-Uhlenbeck (Vasicek) simulation

\[ dr_t = \kappa(\theta - r_t) \, dt + \sigma \, dW_t \]
Simulation of general diffusion processes

\[ dr_t = \kappa (\theta - r_t) \, dt + \sigma \, dW_t \]

<table>
<thead>
<tr>
<th>Model</th>
<th>Specification</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \gamma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merton</td>
<td>( dr_t = \alpha dt + \sigma dW_t )</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Vasicek</td>
<td>( dr_t = (\alpha + \beta r_t) dt + \sigma r_t^{1/2} dW_t )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CIR SR</td>
<td>( dr_t = (\alpha + \beta r_t) dt + \sigma r_t^{1/2} dW_t )</td>
<td>0</td>
<td>0</td>
<td>1/2</td>
</tr>
<tr>
<td>Dothan</td>
<td>( dr_t = \sigma dW_t )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GBM</td>
<td>( dr_t = \beta r_t dt + \sigma r_t dW_t )</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Brennan-Schwartz</td>
<td>( dr_t = (\alpha + \beta r) dt + \sigma r_t dW_t )</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CIR VR</td>
<td>( dr_t = \sigma r_t^{3/2} dW_t )</td>
<td>0</td>
<td>0</td>
<td>3/2</td>
</tr>
<tr>
<td>CEV</td>
<td>( dr_t = \beta r_t dt + \sigma r_t^{\gamma} dW_t )</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
END
Appendix: Functionalities Overview

FinMetrics: Main Functionalities I

• **Time series manipulation**
  + Merging, aggregation, lags, returns, missing values,...
  + graphics

• **Univariate time series**
  + acf, pacf, normality-, serial corr. Tests
  + spectral analysis
  + ARIMA, AIC, BIC
  + prediction, simulation

• **Vector AR models (VAR)**
  + Estimation, inference, model selection
  + Forecasting & simulation
  + Structural analysis
  + Bayesian VAR

• **Unit root tests**
  + ADF, PP
  + KPSS for stationarity
  + efficient unit root tests: ERS, DF-GLS, modified PP

• **Cointegration**
  + Phillips-Ouliaris distr.
  + Johansen`s LR tests
  + ML estimation of VECM

• **Long memory times series**
  + R/S, GPH-test, periodogram-, Whittle`s method
  + FARIMA, SEMIFAR
  + FIGARCH, FIEGARCH
  + prediction, simulation

• **Rolling analysis**
  + rolling analysis of general models
  + technical analysis indicators
  + backtesting

• **Robust change detection**
  + REGARIMA
  + Robust REGARIMA
Appendix: Functionalities Overview

Main Functionalities II

• **Univariate GARCH**
  - ARCH, GARCH
  - EGARCH, TGARCH, PGARCH, GARCH-M
  - exog. variables in conditional mean or variance eq.
  - non-Gaussian error (t-distr., GED)
  - fitting, prediction, simulation

• **Multivariate GARCH**
  - EWMA, DVEC
  - matrix diagonal, BEKK, CCC, PC
  - exogenous variables and arma terms
  - conditionally t-distr. residuals
  - diagnostics (JB, SW, LB, LM)
  - prediction & simulation Modelling extreme values

• **Time series regression**
  - OLS
  - Eiker-White HC and Breusch-Pagan- & White-tests
  - Newey-West HAC:
    - Recursive least squares

• **Systems of regression equations**
  - SUR, FGLS, ML
  - Nonlinear SUR

• **State Space Models**
  - Kalman filter, -smoother
  - Forecasting, simulation
  - ML est. Of unknown param., Quasi ML
Main Functionalities III

• **Modelling extreme values**
  + Extremes over threshold (POT)
  + Hill’s estimator

• **Copulas**
  + Parametric copula classes and families
  + Fitting and simulation

• **Factor models for asset returns**
  + Statistical factor models for returns & principal components
  + Macroeconomic- and fundamental factor models

• **Term structure of interest rates**
  + Rate conversions (discount, spot and forward rate)
  + Fitting term structure by
    • Regression splines, smoothing splines
    • Parametrics models: (Nelson-Siegel, -Svensson)

• **Nonlinear time series models**
  + BDS test for nonlinearity
  + TAR, SETAR, STAR
  + Markov switching state space models

• **Generalized methods of moments**
  + GMM estimation, hyp. testing for lin. Models

• **Continuous time models in finance**
  + Approx. Solutions to SDE
  + Simulators

• **Semi nonparam. conditional density**
  + SNP models, selection, diagnostics, prediction

• **Efficient methods of moments**
• primal-dual interior point method with higher order correction for Linear Programming (LP) models. \{higher\}
• simplex method for Linear Programming (LP) and mixed integer programming (MIP) models. \{simplex\}
• primal-dual interior point method based on line search for general Convex Programming (CP) models. \{line\}
• primal-dual interior point method based on trust region method for general Non-Linear Programming (NLP) models. \{trust\}
• primal-dual interior point method based on quasi-Newton method for general Non-Linear Programming (NLP) models. \{BFGS\}

<table>
<thead>
<tr>
<th></th>
<th>LP</th>
<th>MIP</th>
<th>CP</th>
<th>NLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>simplex</td>
<td>best fit</td>
<td>base fit‡</td>
<td></td>
<td></td>
</tr>
<tr>
<td>higher order</td>
<td>base fit‡</td>
<td>base fit‡</td>
<td></td>
<td></td>
</tr>
<tr>
<td>line search</td>
<td>fit</td>
<td></td>
<td>base fit</td>
<td></td>
</tr>
<tr>
<td>trust region</td>
<td>fit‡</td>
<td></td>
<td>base fit‡</td>
<td></td>
</tr>
<tr>
<td>BFGS</td>
<td>fit</td>
<td></td>
<td>best fit</td>
<td></td>
</tr>
</tbody>
</table>

‡ The default in NUOPT
Not applicable

Table 1: Types of problems (across the columns) and the optimization algorithms in NUOPT (along the rows).