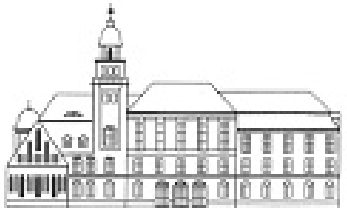
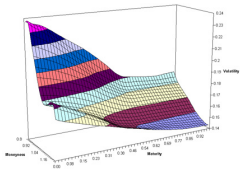
MD*ReX
Modern Statistical Tools in Office Application

Gökhan Aydınlı
Wolfgang Härdle

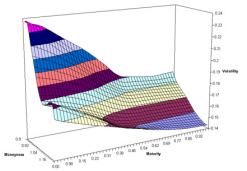
Sonderforschungsbereich 373
Center for Applied Statistics and Economics (CASE)





Outline of the talk

1. Introduction
2. Review
3. Philosophy
4. Technical Implementation
5. Illustration
6. Conclusion
7. References



Introduction
Why Excel?

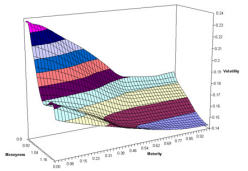
Excel is an highly recognized and industry wide used spreadsheet application

Excel offers high usability and many features: various layout facilities for presentation purposes, formulas for transformations and calculations, sorting, filtering...

Excel however provides only basic statistical skills

Statistical procedures available in Excel are performing poorly (Knüsel (1998) and McCullough and Wilson (1999))

Sophisticated statistical methods arent available (e.g. Panel Data Analysis, Linear Regression with Heteroskedasticity,...)

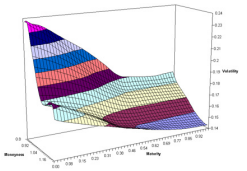


Why XploRe?

XploRe is a high-performance statistical environment, especially for large scale statistical analyses either in business, science or education offering a procedural programing language...

...but XploRe is not a spreadsheet application!

$\Rightarrow$ It is an opportunity to enable Excel to carry out high level statistical procedures by integrating it with statistical software

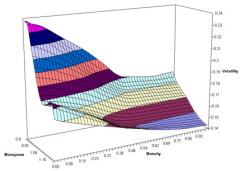


Why MD*ReX?

MD*ReX combines Excel and XploRe: MD*ReX fills the gap between the every day work with office applications and the need for accurate, powerful and reliable computational statistics

MD*ReX is nested within a well known environment, has the same look and feel as MS Office applications and seamlessly fits into Excel

Enables the user to combine the powerful graphical engine of Excel with the high-grade statistics skills of XploRe



Alternatives?

Add-ins (*.xla)

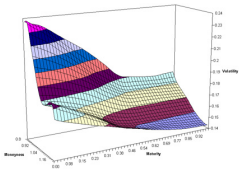
- xlStat
- xlStatistics
- Unistat
- Analyse-It

Integrate with available Software via Dynamic Linking

- Matlab
- Statistica

COM based statistical Computing

- R
- XploRe



What is new?

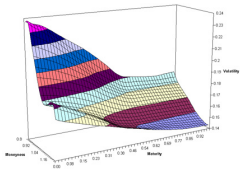
Introducing an alternative solution by integrating the statistical computing environment XploRe and Excel

Architecture provides local computing intelligence as well as a true client / server setup

Java based middle-ware technology

Support for distributed business models e.g.:

ebook Springer



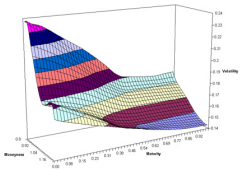
Review
XploRe

Combination of classical and modern statistical procedures

Basis for statistical analysis, research and teaching

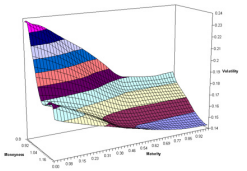
Purpose lies in the exploration and data analysis

Provides statistical engine services via XQS

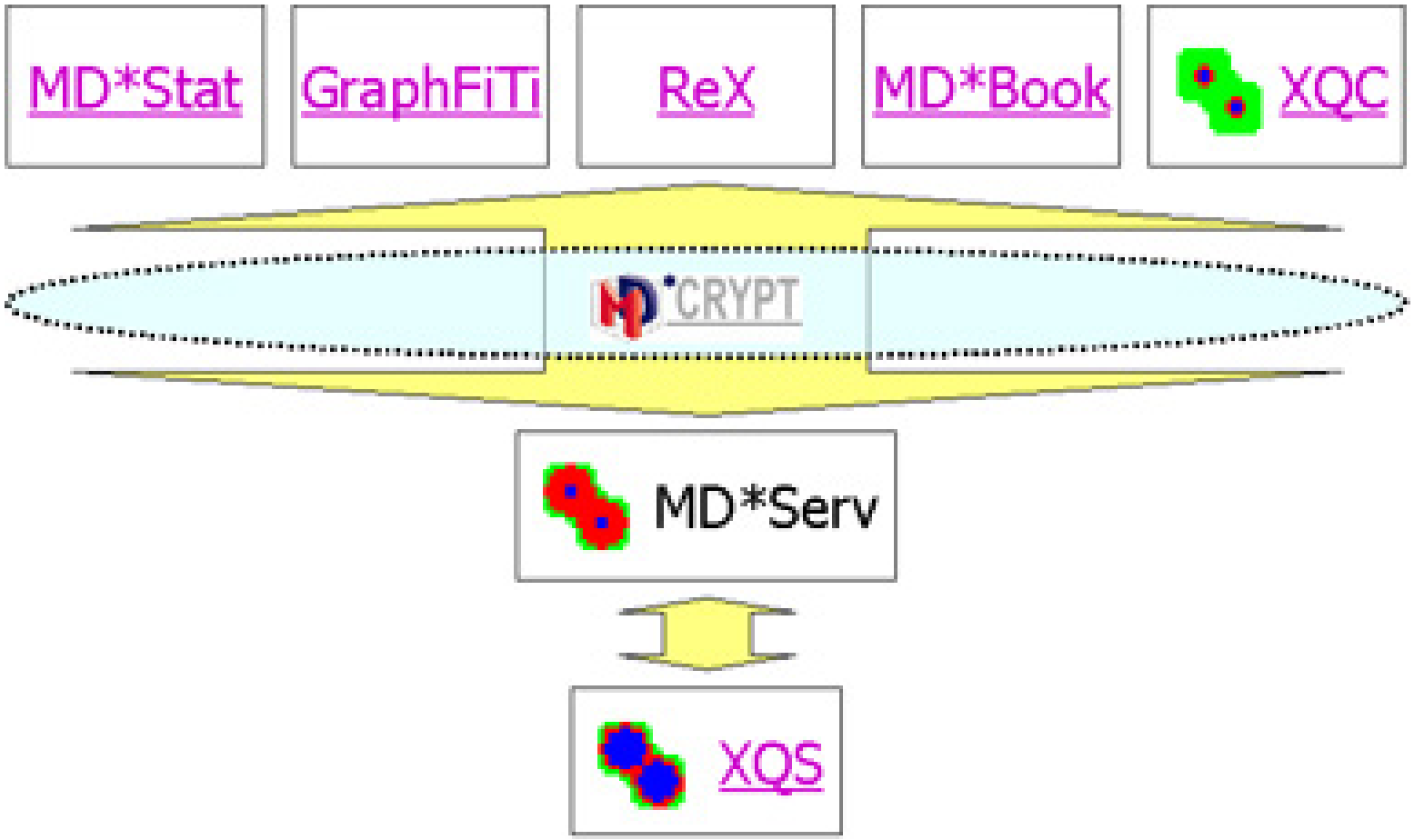


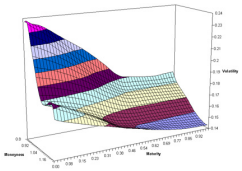
Product Portfolio for XploRe

-  XQS (XploRe Quantlet Server)
-  XQC (XploRe Java Quantlet Client)
-  MD*Crypt
-  Quantlets
-  MM*Stat
-  MD*Book



Techniques behind XploRe

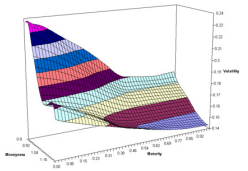




Excel

+ A widely used spreadsheet application
+ Flexibility: manipulates data interactively
+ Handling / Usage: intuitive and easy to use GUI, wide spectrum of built-in features, functions and wizards for data analysis and presentation
+ Powerful graphic engine

- Does not provide further statistical intelligence (e.g. Wavelets, Neural Networks, Time series analysis)
- Inaccurate statistical capabilities



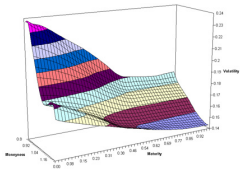
Philosophy of MD*ReX

- Offers various ready to use statistical procedures
- Local XQS is triggered automatically when MD*ReX is started
- Built-in networking functionality within a C/S framework
- Operates any numerical XploRe functions from Excel
- Three layer philosophy:

Frontend: Excel / MD*ReX

Middleware: MD*Serv / MD*Crypt

Backend: XQS

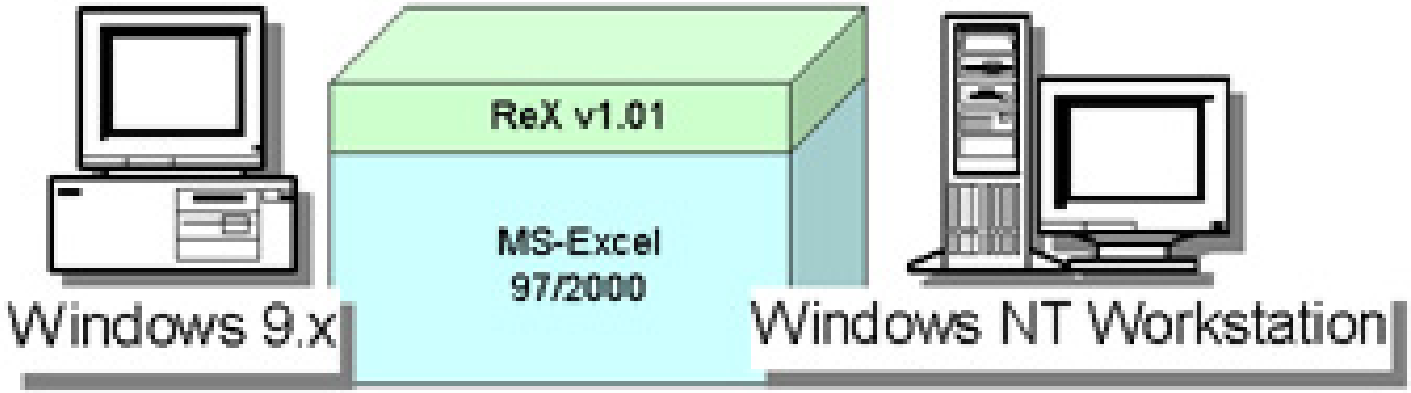


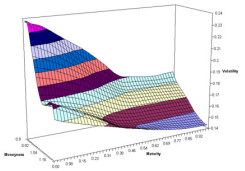
Technical Aspects
Basic Architecture

MD*ReX is a COM (Component Object Model) Add-In for Microsofts

Office application

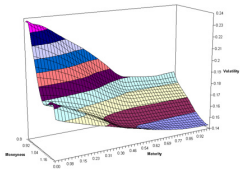
Applicable Operating Systems (OS) are Microsoft Windows 95/98/2000/NT





Software Requirements

- Microsoft Windows9.x, Windows NT Workstation (with SP5 or higher)
- Microsoft Excel 8.0 (97) or Excel 9.0 (2000)
- MD*ReX
- MD*Crypt (mandatory)
- MD*Serv (optional)
- XQS XploRe Quantlet Server (optional)
- Java environment (Microsoft and SUN, mandatory)

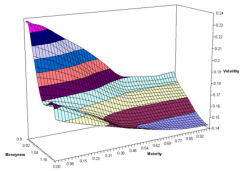


Architecture

MD*ReX operates as an intermediary 32-bit client application between Excel and XQS

MD*ReX is written in Visual Basic and implemented as COM DLL

Communication to the XQS is accomplished through a custom COM version of MD*Crypt: cryptXP.dll

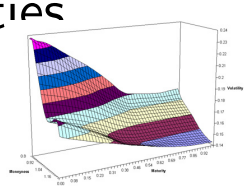


Schematic Architecture

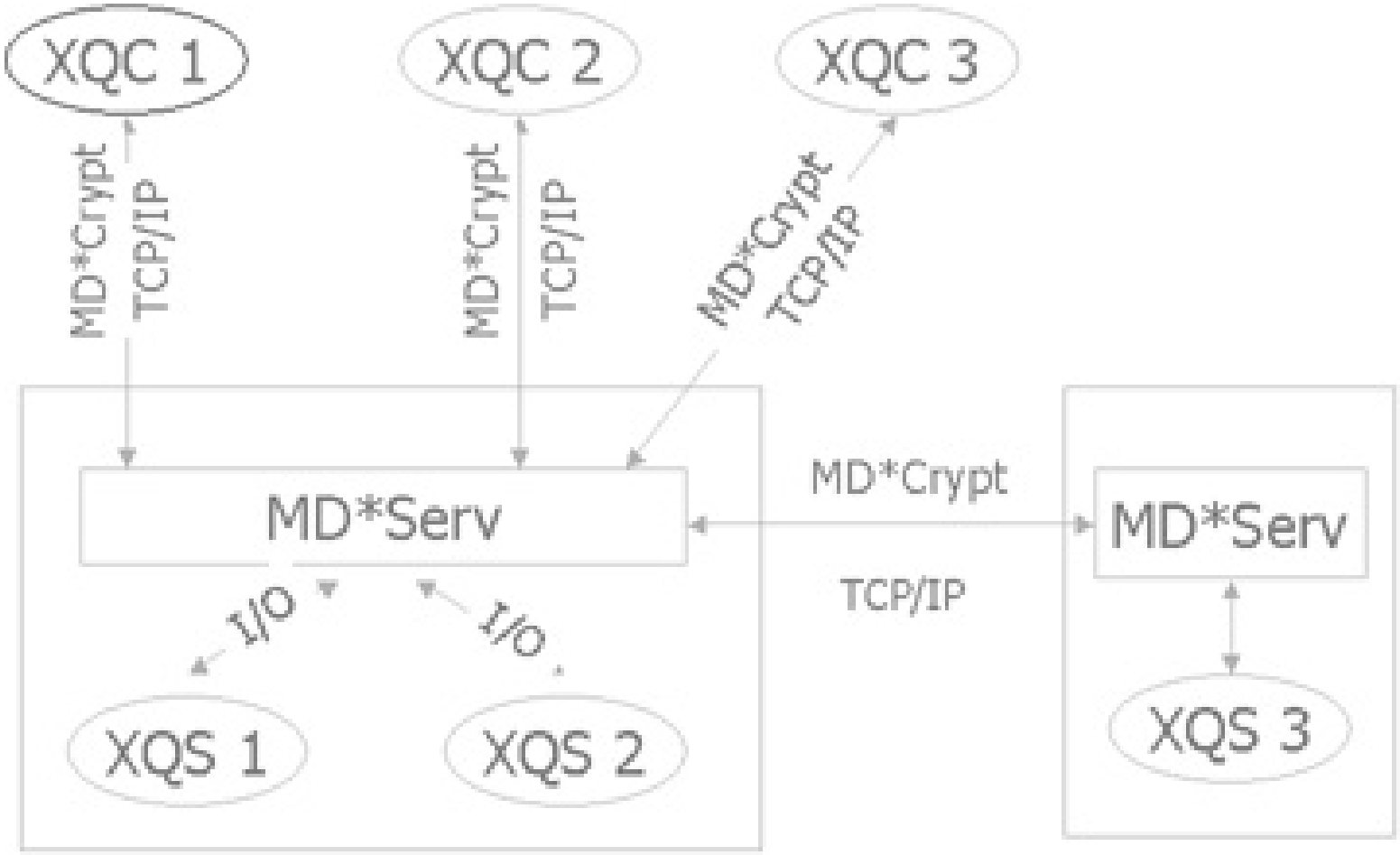
MS Excel provides intuitive spreadsheet functionality

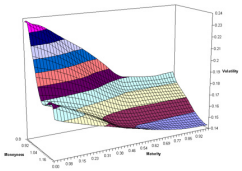
MD*ReX runs within MS Excel

The XQS provides the connection to XploRes computational capabilities

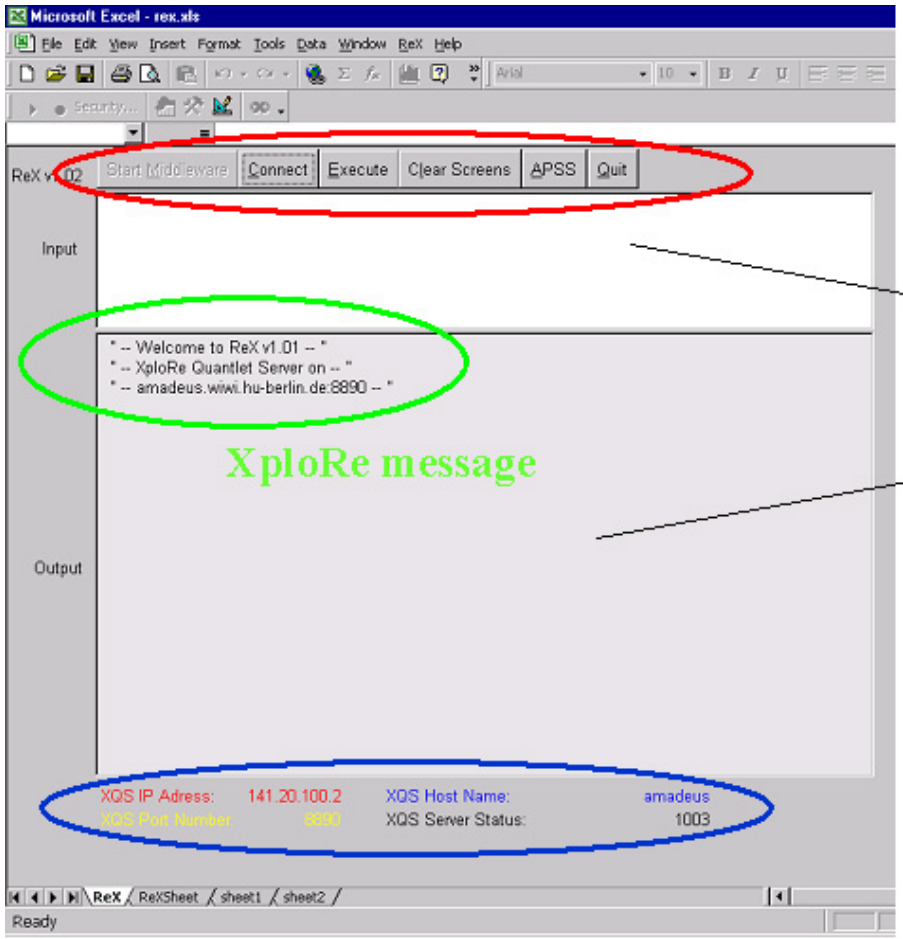


XploRe's Client/Server Architecture





MD*ReX from the User's Perspective I

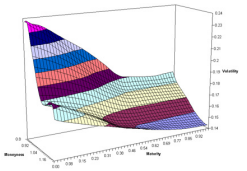


ReX command buttons

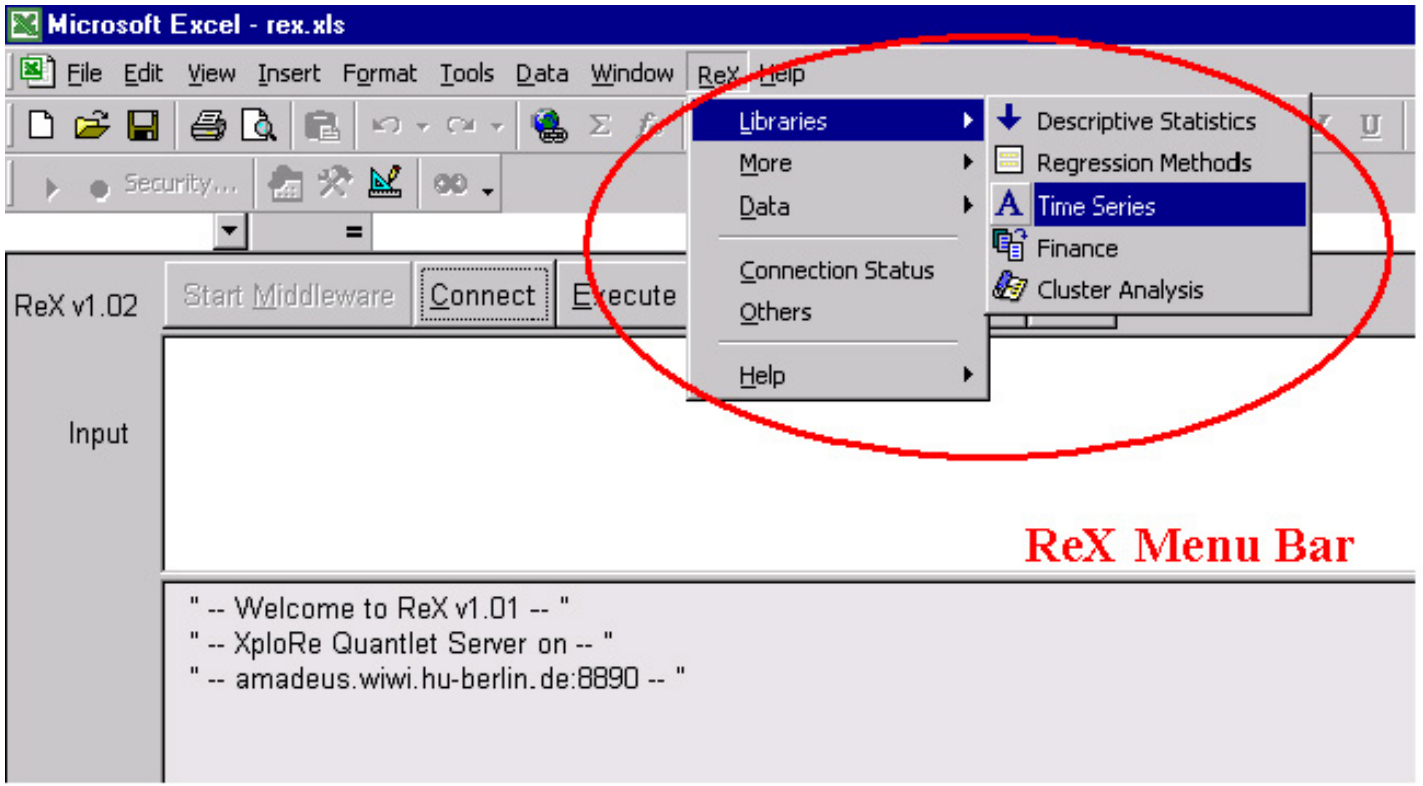
XploRe Direct Editor

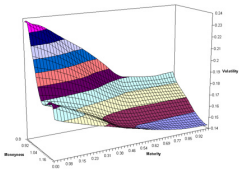
Output Window

ReX status bar

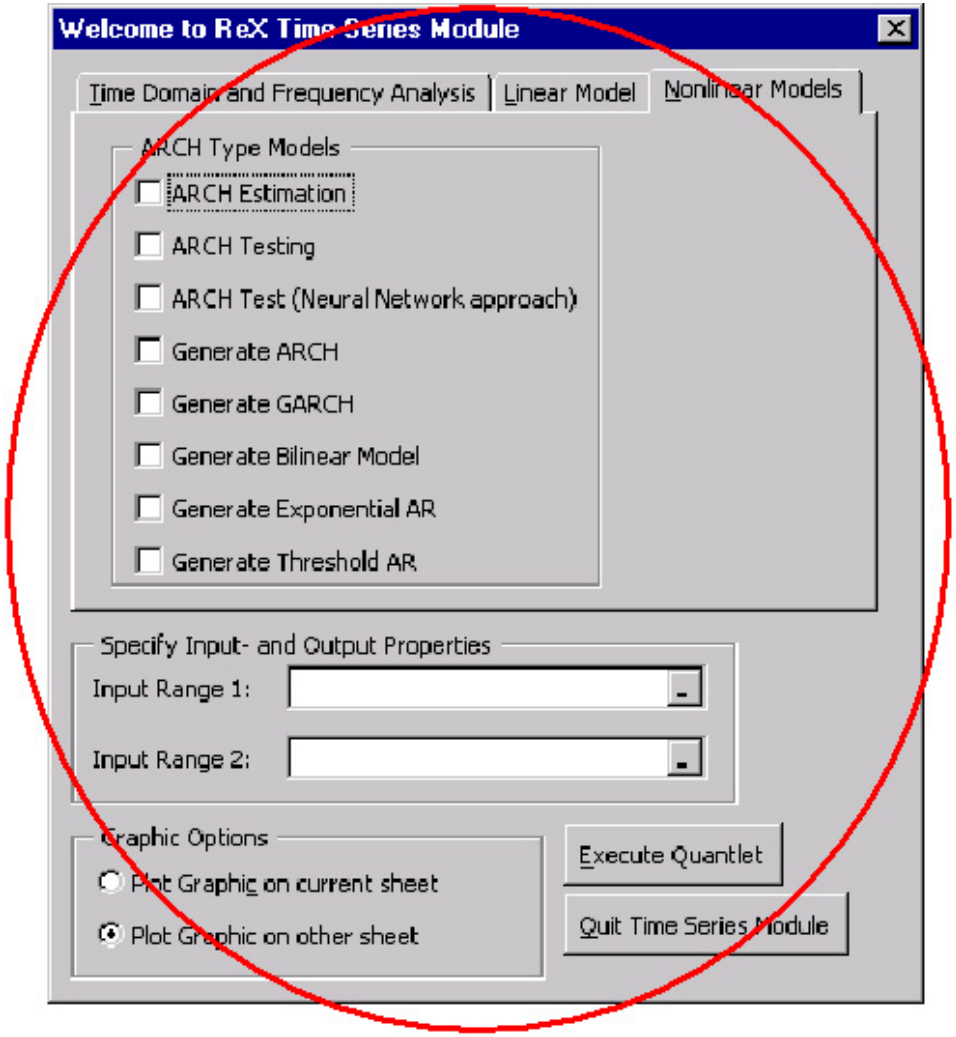


MD*ReX from the User's Perspective II

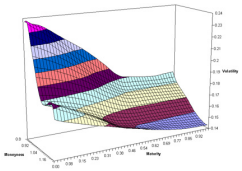




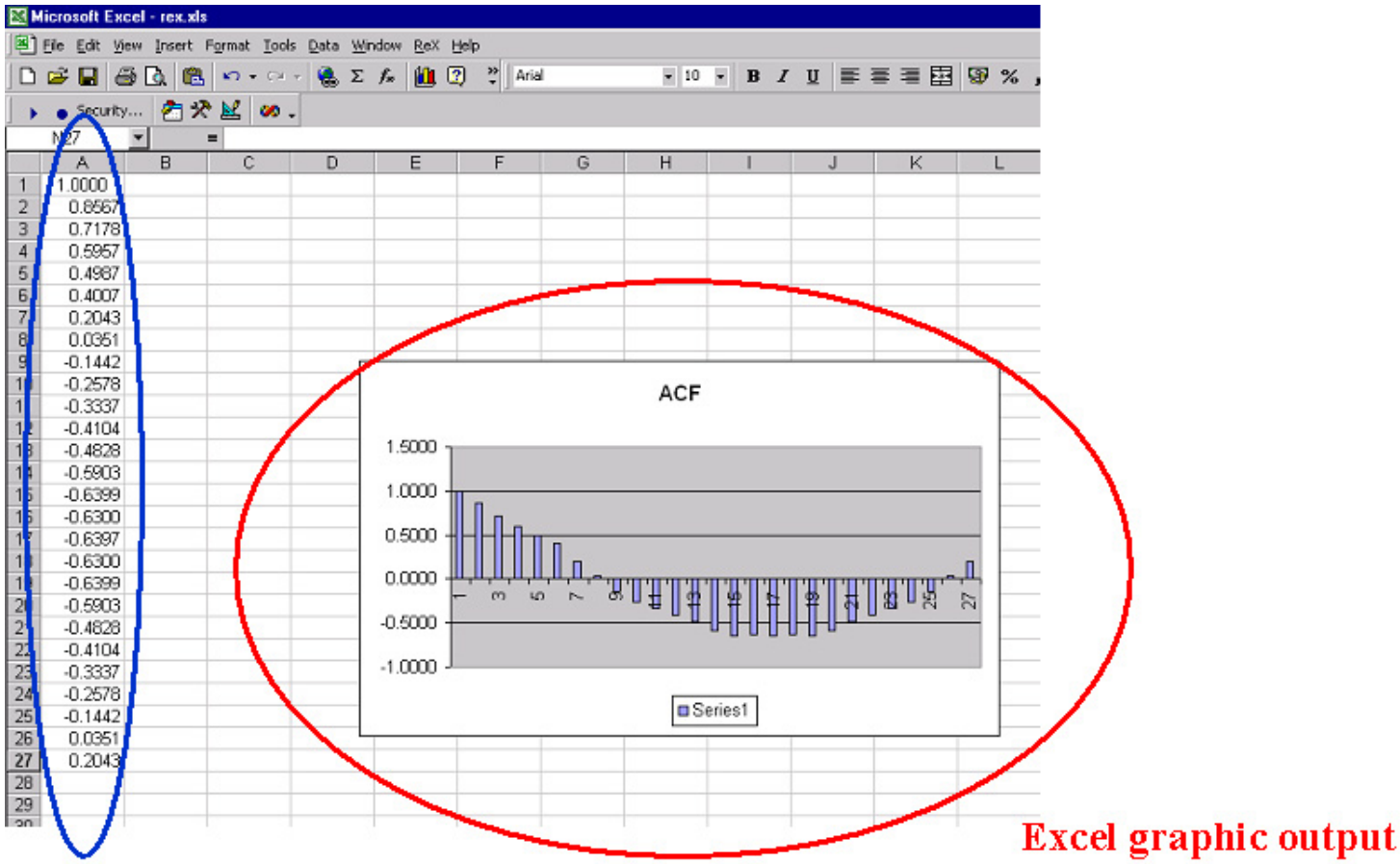
MD*ReX from the User's Perspective III



ReX Time Series Module

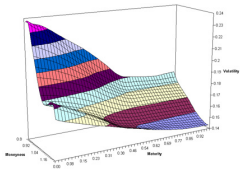


MD*ReX from the User's Perspective IV

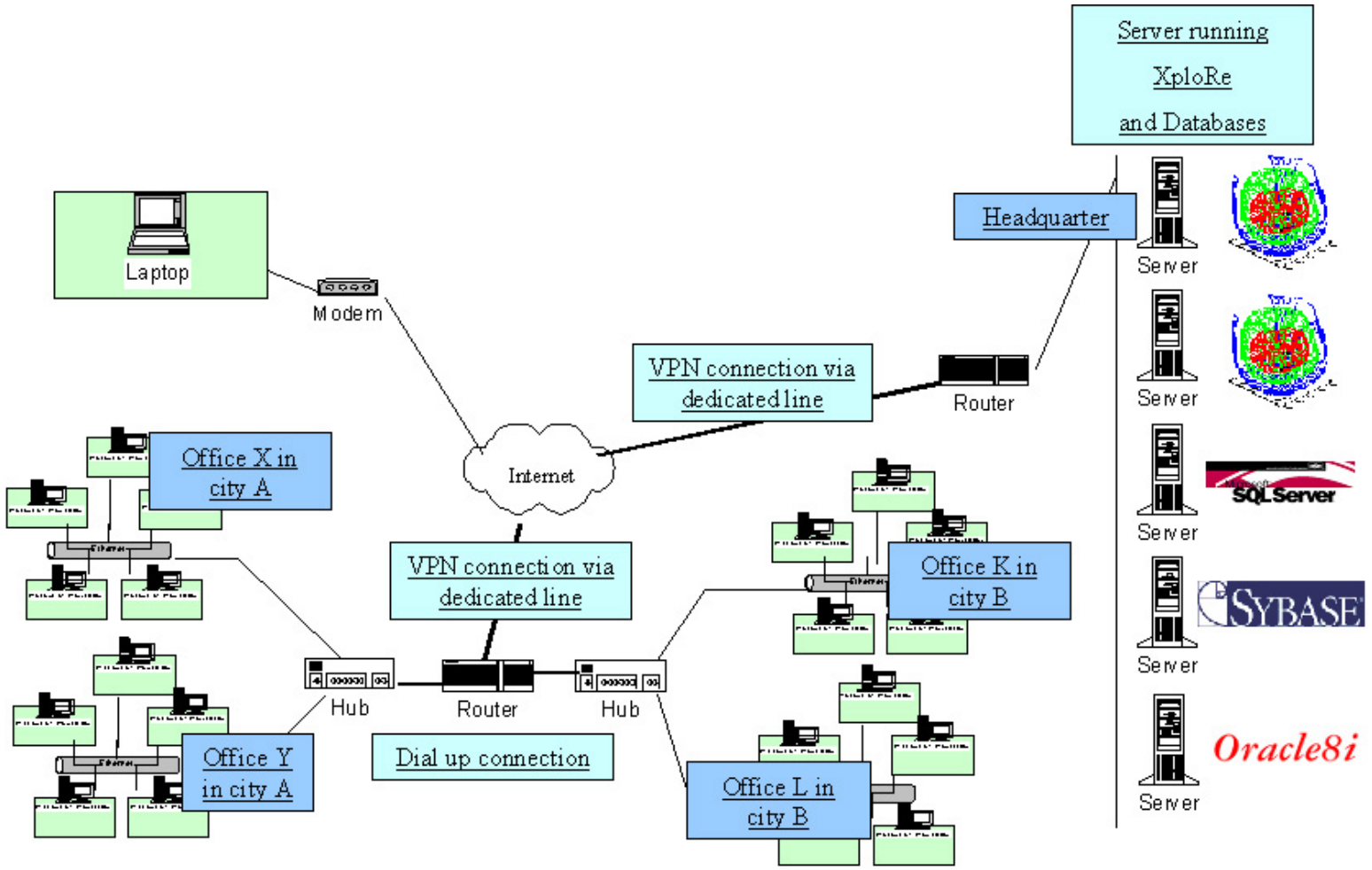


XploRe numeric output

Excel graphic output

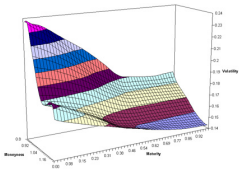


MD*ReX in Heterogenous Environments



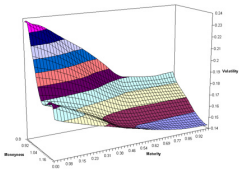
Concluding Remarks

- It is possible to integrate Excel with the statistical environment XploRe
- It enables user to access XploRe under Office applications
- Some sophisticated statistical procedures can be conducted in e.g. Excel and presented in Word



Future Development

- Customize existing libraries of MD*ReX to function properly with MD*ReX
- Complete merger of XploRe functionality into MD*ReX
- Development of GUI and usability



References

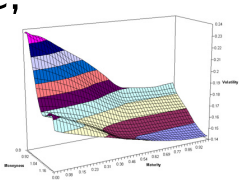
Aydınlı, G. (2001). *Net Based Spreadsheets in Quantitative Finance*, in W. Härdle, T. Kleinow, G. Stahl: Applied Quantitative Finance, Springer Verlag, Heidelberg.

Aydınlı, G., Härdle, W., Kleinow, T. and Sofyan, H.(2001). *ReX: Linking XploRe to Excel*, ISM Symposium "Statistical Software in the Internet Age", February 19-21, 2001.

Härdle, W., Hlavka, Z. and Klinke, S. (2000). *XploRe Application Guide*, Springer Verlag, Heidelberg.

Härdle, W., Kleinow, T. and Tschernig, R. (2000). *Web Quantlets for Time Series Analysis*, SFB-373 Discussion Paper, Humboldt-Universitä zu Berlin.

Härdle, W., Klinke, S. and Müller, M. (1999). *XploRe Learning Guide*,



Springer Verlag, Heidelberg.

Knüsel, L. (1998). *On the Accuracy of Statistical Distributions in Microsoft Excel 97*,The Statistical Software Newsletter in Computational Statistics & Data Analysis, 26, 375-379.

McCullough, B.D. and Wilson, B. (1999). *On the Accuracy of Statistical Procedures in Microsoft Excel*, Computational Statistics & Data Analysis, Vol. 31, p. 27-37.

Neuwirth, E. and Baier, T. (2001). *Embedding R in Standard Software, and the other way round*, DSC 2001 Proceedings of the 2nd International Workshop on Distributed Statistical Computing.