Recent Advances and Trends in Time Series Analysis: Nonlinear Time Series, High Dimensional Inference and Beyond

Yulia Gel (University of Waterloo),
Rafał Kulik (University of Ottawa),
Hao Yu (University of Western Ontario)

April 27, 2014–May 2, 2014

1 Overview of the Field and Recent Developments

With the advent of massive and complex data sets and especially phenomenon of high frequency and ultra high frequency observations, the classical paradigms of linear time series and the related Box-Jenkins methodology is not applicable, which leads to a booming interest in developing novel methodology in time series analysis. In particular, it is now well recognized that many financial, biological and environmental observations as well data from complex dynamic computer networks exhibit many similar patterns such as *heavy tails*, *clustering of extremes*, *weak and long range dependence*, and *non-stationary behavior*, which are not well described by the linear models. Those empirical observations led to development of non-linear models such as GARCH (Robert Engle was awarded the Noble Prize in Economics in 2003 for its invention) or stochastic volatility (both in a univariate and multivariate setting), models with local stationarity and/or time varying coefficients, or Poisson shot-noise processes, the latter especially in a context of network traffic modeling. However, while probabilistic structure for some of the models is reasonably well-understood, statistical inference and, especially, its extension to multivariate non-linear time series with the above mentioned characteristics are yet not well developed and constitute a challenging issue both from theoretical and applied point of view.

Moreover, many modern time series methods require transcending traditional boundaries between time series analysis, nonparametric and spatial statistics, applied probability, and discrete mathematics, which leads to fusion of knowledge and ideas across various branches of statistics and mathematics. Such examples include functional data analysis methodology employed to describe time-space phenomena, the recent advances on high dimensional inference of ultra high frequency financial time series and characterizing dynamics of time series as a random network process.

It is also well recognized that theoretical and methodological advances in time series analysis are possible only if the gaps between paradigms of various branches of statistics and applied sciences are bridged together, forging joint interdisciplinary efforts. The workshop aimed to take advantage from the state-of-the-art research from multiple disciplines and builds on the fusion of complementary expertise of leading researchers in time series, nonparametrics, multivariate analysis, functional data analysis, applied probability and computational statistics.

Two main themes of the workshop were **Nonlinear Time Series** and **High Dimensional Inference** applied in the time series context. The list of topics includes

1. High dimensional, multivariate time series;
2. Short memory processes;
3. Long memory processes;
4. Non-stationary models;
5. Non-linear time series;
6. Heavy tailed time series;
7. Change-point and trend detection problems;

The speakers presented variety of methods to deal with non-linear, high dimensional time series, starting with classical methods, and including very recent techniques such as:

- adaptive estimation (wavelets, spectral analysis);
- Bayesian techniques;
- functional data analysis methods;
- bootstrap and resampling techniques;
- empirical likelihood.

2 Presentation Highlights

Monday, April 28: The theme for the morning session was *long memory processes*. The conference was opened by talks of **Peter Robinson** and **Liudas Giraitis**. The first speaker focused on more applied aspects of estimation of long memory parameter for panel data, while the second speaker talked about new results on existence of second-order stationary ARCH(∞) processes, that include FIGARCH or IGARCH families. This was a long-standing problem in the area of nonlinear time series with long memory (see [8]).

The second session of the day featured **Ejaz Ahmed** and **Daniel Pena**, who gave talks on statistical aspects of *high dimensional*, possibly *non-stationary time series*. The first speaker presented a new high-dimensional shrinkage estimation strategy ([1]), while the second speaker proposed a new time domain procedure to define dynamic principal components (DPC) that is applicable to non-stationary and relatively short series (([15])).

In the afternoon sessions we continued with statistical aspects of *non-stationary time series*: **Rainer Dahlhaus** discussed local polynomial fitting of time-varying parameter curves ([5]), **Eunice Menezes** presented an overview of wavelets methods, while **David Stoffer** discussed Bayesian methodology ([18]). **Piotr Kokoszka** discussed functional data approach (see [12]) to financial time series, with particular applications to intraday price.

Tuesday, April 29: As on Monday, we started with *long memory processes*. **Murad Taqqu** considered vector-valued multilinear polynomial-form processes with either short or long memory components and discussed possible limiting distributions. **Vladas Pipiras** continued with multivariate long range dependence discussing possible definitions and new parameters modeling the cross spectrum.

In the next session the speakers focused on *dependent random fields*. **Gail Ivanoff** presented a new martingale technique to deal with limit theorems for short range dependent causal linear process, while **Jens-Peter Kreiss** considered more statistical issues of bootstrap for random fields.

The first afternoon session featured **Francois Roueff**, **Mohsen Pourahmadi** and **Nozer Singpurwalla**. The first speaker showed how to use multivariate point processes (more specifically, Hawkes processes) for modeling limit order book of a financial asset. Pourahmadi discussed thresholded generalized principal

component regression in the context of multivariate time series ([16]). The session concluded with a more philosophical discussion on connection of quantum physics, probability foundation and statistical inference by Singpurwalla ([13]).

The last session of the day dealt with *heavy tails and point processes*. **Zhengyan Lin** discussed weak convergence of various general functionals of partial sums of dependent random variables statistics to stochastic integrals driven by either Brownian motion or Levy-stable processes. **Bojan Basrak** continued with functional limit theorems for heavy tailed time series, using a non-standard topology (see [3]). Finally, **Robert Lund** showed how to simulate stationary count time series, with a particular focus on a much less investigated phenomenon of negative correlation ([7]).

Wednesday, April 30: The topic of the first session was *high-dimensional autocovariance matrices*. **Thomas Mikosch** presented limit theorems for the largest eigenvalues in a sample covariance matrix for multivariate time series with heavy tails (see [6]). The next speaker, **Timothy McMurry** focused on light-tailed time series and a classical problem of linear prediction, however in a high-dimensional setting.

The second session featured two talks on *change-point problems*. **Herold Dehling** presented his work on robust change-point tests for stationary time series with both short- and long memory, using two-sample U-processes and U-quantile processes, while **Michael Baron** discussed a Bayesian approach to change-point detection in multivariate time series.

The afternoon was free. The participants enjoyed a beautiful visiting Banff and its neighbourhood or Lake Louise.

Thursday, May 1: The first session featured two talks on *time series extremes*. **Holger Drees** presented his new work on estimation of the distribution of a tail chain that appears as a weak limit of a heavy tailed time series, conditionally on one component being large (see [9]). **Gemal Chen** discussed some results obtained in studying finite sample dependent extremes.

The topic of the next session was *limit theorems for weakly dependent time series*. **Zhou Zhou** discussed central and noncentral limit theorems for weighted V-statistics in case of nonstationary nonlinear processes, while **Martin Wendler** considered bootstrap for weakly dependent Hilbert space-valued random variables.

During the third session the speakers presented recent work on different statistical issues for dependent time series data in *spectral domain*. **Dan Nordman** discussed a frequency domain empirical likelihood method for irregularly spaced data ([2]). The main problem that arises is that the lack of the usual orthogonality properties of the discrete Fourier transform for irregularly spaced data. **Piotr Fryźlewicz** also focused on a spectral domain approach to time series, introducing a hierarchically-ordered oscillatory basis of simple piecewise-constant functions that allow for detection of change-points in high frequency data ([10]). Finally, **Sofia Olhede** discussed Whittle likelihood method for nonstationary multivariate processes ([20]).

The last session of the day featured **Lilia Leticia Ramirez Ramirez** and **Slava Lyubchich** who discussed *statistical issues for multivariate time series*. The first speaker presented a method for trend estimation of multivariate time series, without relaying without relying on specific models for the trend and noise components ([17]). The method was applied to the Mexican macroeconomic data. Lyubchich discussed nonparametric methods for nonlinear trend detection and trend synchronism in multiple time series using local factor approach and hybrid bootstrap ([14]).

Friday, May 2: The final morning featured three speakers. **Edit Gombay** discussed change detection for time series following Generalized Linear Models suing likelihood methods. **Rogemar Mamon** used multivariate Ornstein-Uhlenbeck processes to forecasting The workshop concluded with **Reg Kulperger** discussing new theoretical results for estimation in GARCH-in-mean processes.

3 Outcome of the Meeting

The main goal of the meeting was to bring researchers working in a very broad field of time series analysis. The participants had an unique opportunity to learn about variety of topics and methods and techniques, as evidenced above. Variety of problems tackles at the meeting is illustrated by the areas of applications con-

sidered at the meeting: finance and economics, climate and weather modeling, flood control, brain signal modeling, clinical trials among others.

References

[1] E. Ahmed, *Penalty, Shrinkage and Pretest Strategies: Variable Selection and Estimation*, Springer Briefs in Statistics, 2014.

[2] S. Bandyopadhyay, S. N. Lahiri, D. J. Nordman, A functional limit theorem for dependent sequences with infinite variance stable limits. *Preprint*, (2014).

[3] B. Basrak, D. Krizmanic, J. Segers, A functional limit theorem for dependent sequences with infinite variance stable limits. *Ann. Probab.* **40** (2012), 2008–2033.

[4] J. Beran, Y. Feng, S. Ghosh, R. Kulik, *Long-memory Processes: Probabilistic Properties and Statistical Methods*, Springer Series in Statistics, (2013).

[5] R. Dahlhaus, Locally stationary processes. *Handbook of Statistics*, 30 (2012).

[6] R. Davis, T. Mikosch, O. Pfaffel, Convergence of the largest eigenvalues of a sample covariance matrix for heavy-tailed multivariate time series. *Preprint*, (2013).

[7] R. Davis, S. Holan, R.B. Lund, N. Ravishanker, *Handbook on Discrete Valued Count Series*, (2014+).

[8] R. Douc, F. Roueff, P. Soulier, On the existence of some ARCH(∞) processes. *Stochastic Processes and Their Applications* **118** (2008), 755–761.

[9] H. Drees and H. Rootzen, Limit theorems for empirical processes of cluster functionals. *Ann. Stat.* **38** (2010), 2145–2186.

[10] P. Fryźlewicz, A. L. Schroeder, Adaptive trend estimation in nancial time series via multiscale change-point-induced basis recovery. *Statistics and Its Interface*, **6** (2013), 449–461.

[11] L. Giraitis, H. L. Koul, D. Surgailis, *Large Sample Inference for Long Memory Processes*. World Scientific Publishing Company Inc., (2012).

[12] L. Horvath, P. Kokoszka, *Inference for Functional Data with Applications*, Springer Series in Statistics, (2012).

[13] J. Landon, F. X. Lee and N. D. Singpurwalla, A Problem in Particle Physics and Its Bayesian Analysis. *Statistical Science* **26** (2011), 352–368.

[14] V. Lyubchich, Y. R. Gel, Detecting Trend Synchronism in Multiple Time Series. *Preprint*, (2014).

[15] D. Peña, V. J. Yohai, Dynamic Principal Components in the Time Domain. *Preprint*, (2014).

[16] M. Pourahmadi, *High-Dimensional Covariance Estimation*, Wiley, 2013.

[17] L.L. Ramirez Ramirez, V. Guerrero Guzman, A. Islas Camargo, Trend estimation of multivariate time series with controlled smoothness. *Preprint*, (2014).

[18] O. Rosen, S. Wood, D. S. Stoffer, AdaptSPEC: Adaptive Spectral Estimation for Nonstationary Time Series. *Journal of American Statistical Association* **107** (2012), 1575–1589.

[19] P. M. Robinson and C. Velasco, Efficient Inference on Fractionally Integrated Panel Data Models with Fixed Effects. *Preprint*, (2013).

[20] A.M. Sykulski, S.C. Olhede, J.M. Lilly, J.J. Early, The Whittle Likelihood for Complex-Valued Time Series. *Preprint*, (2014).