

# GEORG M. GOERG

New York, NY 10001, USA  
January 23, 2015

[www.gmge.org](http://www.gmge.org)  
[im@gmge.org](mailto:im@gmge.org)

## PROFESSIONAL EXPERIENCE

### **Google, Inc.**

**New York City, USA**  
**January 2013 – Present**

*Statistician, Quant Analyst*

- member of the quantitative marketing (QM) group
- research, development, and implementation of statistical models in advertising

*Decision Support Analyst Intern*

**June – August 2011**

- intern in the quantitative marketing (QM) group: i) developed and implemented algorithms for change-point detection in time series; ii) research on search query classification and its possible use for improved advertisement strategies.

### **Pontificia Universidad Católica de Chile**

**Santiago, Chile**

*Researcher and Lecturer, Dept. of Statistics*

**March – December 2008**

- researcher (→ “Publications” and Lambert W random variables) and instructor/TA (→ “Teaching Experience”); mainly collaborated with [Wilfredo Palma](#).

## EDUCATION

### **PhD, Statistics**

**August 2009 – December 2012**

[Carnegie Mellon University \(CMU\)](#), Pittsburgh, USA

- Thesis on “Learning Spatio-Temporal Dynamics: Nonparametric Methods for Optimal Forecasting and Automated Pattern Discovery”
- Advisors: [Cosma Shalizi](#) and [Larry Wasserman](#), Department of Statistics
- ASA award for best JSM 2012 [student paper](#) in “Statistical Learning and Data Mining”
- ASA award for best JSM 2011 [student paper](#) in “Statistical Learning and Data Mining”
- Successfully completed the [Future Faculty Program](#)

[Santa Fe Institute \(SFI\)](#), Santa Fe, USA

**June 2012**

- Participant of the 2012 [Complex Systems Summer School \(CSSS\)](#)

### **Diplom-Ingenieur (MS), Mathematics**

**2002 – 2007**

With distinction at the [Vienna University of Technology \(UT\)](#), Vienna, Austria

- Thesis: *Long Memory versus Structural Breaks: A Time-Varying Memory Approach*
- Advisor: [Manfred Deistler](#), Department of Econometrics and System Theory

[City University of New York \(CUNY\)](#), New York City, USA

**Fall 2006**

Exchange term at Hunter College, CUNY Graduate Center, and City College

- ASA award for best JSM 2007 student paper
- Supervisor: [Dana Sylvan](#), Department of Mathematics & Statistics, Hunter College

## SKILLS

### Languages

German (native), English & Spanish (fluent), Portuguese (intermediate), French (basics), Arabic (beginner), and Latin.

### Programming languages

Programming & Data Analysis: R/S-Plus, Python; SQL; started with Go and MapReduce.  
Typesetting & Web: L<sup>A</sup>T<sub>E</sub>X, knitr / Sweave; XHTML & CSS, CMS (Joomla)  
Math & Others: basics of Matlab, Maple

## SOFTWARE & CODE

### R packages – publicly available at CRAN

Author and maintainer of [LSC](#), [LICORS](#), [LambertW](#), and [ForeCA](#).

Co-author of [afmtools](#)

### Python

PLATO: a statistical analysis tool + GUI for  $N$ -dimensional time series; in particular 3D trajectories of molecular dynamic simulations.

## PUBLICATIONS

### Books

GMG (2010). “Time Series Analysis of Long Memory versus Structural Breaks: A Time-Varying Memory Approach”, Publisher: *Verlag Dr. Müller*. ISIN: 3639246012.

### Peer-reviewed

GMG (2014) “The Lambert Way to Gaussianize skewed, heavy-tailed data with the inverse of Tukey’s h transformation as a special case”. *The Scientific World Journal*, ID 909231, ([www.hindawi.com/journals/tswj/aa/909231/](http://www.hindawi.com/journals/tswj/aa/909231/)).

GMG (2013). “Forecastable Component Analysis”. *JMLR W&CP 28 (2)*: 6472 ([arxiv.org/abs/1205.4591](http://arxiv.org/abs/1205.4591)).

GMG and Shalizi (2013). “Mixed LICORS: A Nonparametric Algorithm for Predictive State Reconstruction”. *AISTATS 2013*: 289-297 ([arxiv.org/abs/1211.3760](http://arxiv.org/abs/1211.3760)).

Althouse, Patterson-Lomba, Hébert-Dufresne, GMG (2013). “The Timing and Targeting of Treatment in Influenza Pandemics Influences the Emergence of Resistance in Structured Populations”. *PLOS Comput Biol* 9(2): e1002912 (1-6).

Patterson-Lomba, Althouse, GMG and Hébert-Dufresne (2013). “Optimizing treatment regimes to hinder antiviral resistance in influenza across time scales” *PLOS ONE* 8(3): e59529 (1-11)

Hébert-Dufresne, Patterson-Lomba, GMG and Althouse (2013). “Pathogen mutation modeled by competition between site and bond percolation”. *Phys. Rev. Lett.* 110, 108103 (1-5)

GMG (2012) “Testing for white noise against locally stationary alternatives”. *Statistical Analysis and Data Mining (SAM)*, 5 (6), p. 478 – 492 ([dx.doi.org/10.1002/sam.11157](http://dx.doi.org/10.1002/sam.11157))

GMG (2011b). “A Frequency Domain EM Algorithm for Time Series Classification with Applications to Spike Sorting and Macro-Economics”. *Statistical Analysis and Data Mining (SAM)*, 4 (6), p. 590 – 603 ([arxiv.org/abs/1103.3300](http://arxiv.org/abs/1103.3300)).

GMG (2011a). “Lambert W Random Variables - A New Generalized Family of Skewed Distributions with Applications to Risk Estimation”, *The Annals of Applied Statistics*, 5 (3), p. 2197 – 2230 ([arxiv.org/abs/0912.4554](http://arxiv.org/abs/0912.4554)).

## Proceedings

GMG and Draghicescu (2007). “Nonparametric modeling of the second order structure of processes with time - varying memory”, 2007 JSM Proceedings, Alexandria, VA.

## Technical Reports

Jin, Koehler, GMG, Remy. “The Optimal Mix of TV and Online Ads to Maximize Reach”. Technical report at Google Inc. ([research.google.com/pubs/pub41669.html](http://research.google.com/pubs/pub41669.html))

## REPRODUCIBLE RESEARCH INTERESTS

Time series, forecasting, long memory, locally stationary, frequency domain methods; non-parametric methods, pattern recognition in space-time systems; skewed and heavy-tailed distributions, Gaussianizing data; Lambert W function; user-friendly statistical computing

## PAPERS IN PREPARATION

GMG and Shalizi. “LICORS: Light Cone Reconstruction of States for Non-parametric Forecasting of Spatio-Temporal Systems”. Submitted ([arxiv.org/abs/1206.2398](http://arxiv.org/abs/1206.2398)).

GMG, Shalizi, Wasserman. “Lebesgue Smoothing: Improving Nonparametric Regression by Averaging Predictions”.

GMG, Kurnikova, Shalizi, Kurnikov. “Extracting chemically useful features from protein folding trajectories”.

GMG, Patterson-Lomba, Hébert-Dufresne, Althouse. “Escaping the poverty trap: modeling the interplay between economic growth and the ecology of infectious disease”. Submitted.

## CONFERENCES & WORKSHOPS

“A Frequency Domain EM Algorithm to Detect Similar Dynamics in Time Series with Applications to Spike Sorting and Macro-Economics”, Invited speaker at *JSM 2011*. Miami, USA. July 30 - August 4.

“The Lambert Way to Gaussianize skewed, heavy-tailed data”. Poster at the *IV Skew Workshop* held in honor of Adelchi Azzalini. Santiago, Chile. May 16 - 19, 2011.

“Lambert W Random Variables - A New Family of Generalized Skewed Distributions”, Speaker at *JSM 2010*. Vancouver, Can. July 30 - August 5.

STEF 2008, Poster at the 1<sup>st</sup> *Symposium on Time Series, Econometrics, and Finance*. Valparaíso, Chile. December 18, 2008.

“Una nueva clase de variables aleatorias asimétricas generalizadas”, Speaker (Spanish) at the 8<sup>th</sup> *Congreso Latinoamericano de Sociedades Estadísticas (CLATSE)*. Montevideo, Uruguay. October 7 - 10, 2008.

“Non-parametric Modeling of Time-varying Long Memory”, Invited speaker at *JSM 2007*. Salt Lake City, USA. July 29 - August 3.

## AWARDS

- Yet again: Awarded Paper by ASA – JSM 2012, winner of the US wide student paper competition on “Statistical Learning and Data Mining” (see Publications).
- Again: Awarded Paper by ASA – JSM 2011, winner of the US wide student paper competition on “Statistical Learning and Data Mining” (see Publications).
- High Potential Program – TUtheTOP, [Vienna University of Technology](http://www.vtu.ac.at)
- Awarded Paper by ASA – JSM 2007 in the Government Statistics, Social Statistics, and Survey Research Methods Section.

TEACHING  
EXPERIENCE

Successfully completed the **Future Faculty Program** at CMU      2009 – 2012

**Carnegie Mellon Qatar, Doha**      May – June 2010

*Summer School:* both courses were equivalent to a full-semester 3hrs class + 1hr lab

- **36-425: An Introduction to Time Series Analysis:** classic methods for time series modeling and applications to real world data
- **70-207: Probability and Statistics for Business Applications:** introductory statistics & probability course with focus on business applications

**Pontificia Universidad Católica de Chile, Santiago**      March – December 2008

Graduate course (entirely in Spanish) at the Department of Statistics

- **Una introducción matemática al análisis de series de tiempo - A Mathematical Introduction to Time Series Analysis.** Designed and taught (in Spanish) my own time series course to graduate statistics students – 4 hrs class plus 2 hrs lab, with course notes (pdf book), mid-terms and final exam, homeworks, final project & presentations.

### Teaching Assistant

**Carnegie Mellon University, Pittsburgh, USA**      August 2009 – May 2012

Office hours, grading & recitations for Computational Finance at Tepper School of Business

- **46-929: Financial time series analysis**
- **46-936: Statistical arbitrage:** implementing trading strategies (in R)
- **46-926: Linear Models/Equity Portfolio Management**
- Several **Introductory Statistics & Probability** courses

**Pontificia Universidad Católica de Chile, Santiago**      March – December 2008

Graduate courses (entirely in Spanish) at the Department of Statistics

- **Advanced Time Series Analysis** (Spanish)
- **Financial Time Series Analysis** (Spanish)

OTHER  
PROFESSIONAL  
EXPERIENCE

**Vienna UT, Dept. of Econometrics and System Theory**      Vienna, Austria

*Project assistant*      January – March 2008

- Set up and estimated models for the *Ageing in Vienna* project: forecasting income distribution of population in Vienna for the coming 30 years.

*Project assistant*      May – June 2006

- Developed macro-economic models for main sectors of the Austrian economy.
- Set up, estimated models from input/output data of Austrian economy; produced forecasts for the main sectors of Austrian economy, which were then subsequently used in an international input/output model (INFORUM EconData project).

REFERENCES

Available upon request.