

SFB 649 Discussion Paper 2012-004

Computational Statistics (Journal)

Wolfgang Karl Härdle*

Yuichi Mori**

Jürgen Symanzik***



* Humboldt-Universität zu Berlin, Germany

** Okayama University of Science, Japan

*** Utah State University, USA

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin



SFB 649 ECONOMIC RISK BERLIN

Article type: Focus Article

Computational Statistics (Journal)¹ CSF0019

Wolfgang Karl Härdle

Correspondence to: haerdle@wiwi.hu-berlin.de

C.A.S.E. Centre for Applied Statistics and Economics, School of Business and Economics, Humboldt–Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

Yuichi Mori

Correspondence to: mori@soci.ous.ac.jp

Dept. of Socio–Information, Faculty of Informatics, Okayama University of Science, 1–1 Ridai–cho, Kita–ku, Okayama 700–0005, Japan

Jürgen Symanzik

Correspondence to: symanzik@math.usu.edu

Utah State University, Department of Mathematics and Statistics, Logan, UT 84322–3900, USA

Abstract

Computational Statistics is an international journal that fosters the publication of applications and methodological research in the field of computational statistics. In this article, we will discuss the motivation, history, some specialties, and the future scope of this journal.

Keywords: International journal, history of the journal, specialties of the journal, online communications

JEL classification: C63, C88, Y30

INTRODUCTION

The focus of papers in the Springer–published journal *Computational Statistics (Comput Stat²)*, accessible at <http://www.springer.com/statistics/journal/180>, is on the contribution to and influence of computing on statistics and vice versa. The journal provides a forum for statisticians, computer scientists, and mathematicians in a variety of fields of statistics such as biometrics, econometrics, data analysis, graphics, simulation, algorithms, knowledge–based systems, and Bayesian computing. In addition to regular research articles, the journal also published hardware, software, plus package reports as well as book reviews since its conception.

¹This activity was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

²Internally, the abbreviation *COST* also has been used in recent years, but this should not be used for any official citation or reference purposes.

In this article, we will briefly look at the predecessor of *Computational Statistics* in the next section. In the following four sections, we will look at the beginning, growth and transitions, specialties, and the future of the journal. A short conclusion wraps up this article.

THE INCUNABULUM

The predecessor of *Computational Statistics* was *Computational Statistics Quarterly (CSQ)*, established in 1984 by Physica-Verlag. After six volumes were published, *CSQ* was completely renewed in 1992 by Wolfgang Karl Härdle in order to make the journal more global by including the outcomes of modern statistics and to establish new paradigms of computational statistics. Modern statistics at that time included non-/semi- parametric models, smoothing techniques, image analysis, pattern recognition, simulation-based inference like a bootstrap, and so on. The journal therefore concentrated on interactive statistics with computers (big workstations at that time) and computer-oriented statistics overcoming traditional approaches in statistics. This restart had to be reflected in a new name: *Computational Statistics*. Hereafter we only focus on the history, some specialties, and the future scope of this new journal, *Comput Stat*.

THE BEGINNING

1992–1998 (Härdle — Scott)

The first issue of the new journal *Computational Statistics* was published (as Vol. 7, No. 1 of the series) by Physica-Verlag, a Springer-Verlag company, in 1992. The original cover design of *Comput Stat* is shown on the left hand side in Figure 1, on which symbolic images of those days were displayed. One of them was a 3D graphic on the display, which was drawn using a SUN workstation by David Scott, based on the Mount St. Helens data in the paper “Smoothing by weighted averaging of rounded points” (Härdle and Scott, *Comput Stat*, 7(2): 97-128, 1992).

Comput Stat was started with two editors, Wolfgang Karl Härdle (Editor-in-Chief, EiC) and David W. Scott, and 19 Associate Editors (AEs). A manuscript submitted to *Comput Stat* by the author(s) was assigned to one of editors, the assigned editor assigned one AE matching the classification of the manuscript, and the AE invited two (or more) referees for the review of the manuscript. The same general review system is still in use today. From 1994 to 1999, Sigbert Klinke helped the editors as Editorial Assistant (EA).

The first issue consisted of two sections, “Original Papers” and “Book Reviews”. A “Software Reviews” section was added in the second issue. Book Review Editors were Michael Schimek (1992-1994), Henning Läuter (1992-1995), Christian Ritter

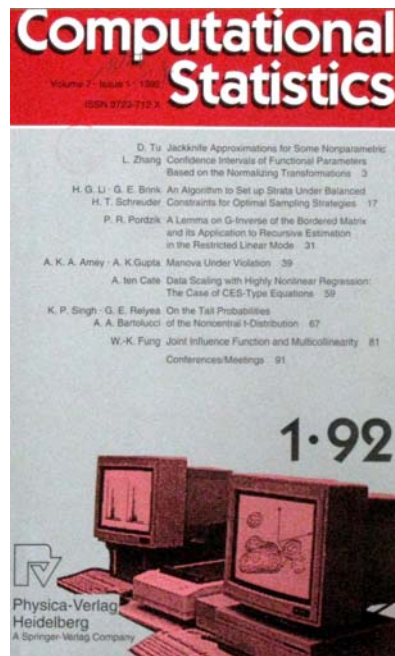


Figure 1: *Comput Stat* cover pages – Old cover page used from 1992 until 2003 (left) and new cover page used since 2004 (right).

(1996-mid 1998), and Andreas Krause (mid 1998-2009). Software Review editor was Rüdiger Ostermann (1992-2010). The journal maintained this three-section structure until 2009/2010 when the Book Review section (in 2009) and the Software Review section (in 2010) were finally dropped.

Although authors were requested to submit three hard copies of their manuscripts by regular mail for about five years after the inception, the way of submission/communication was replaced by e-mail in 1996.

The number of submissions fortunately increased from about 20 in 1994 to about 50 in 1996 as shown in the left chart in Figure 2.

GROWTH AND TRANSITIONS

1999–2003 (Härdle — Newton — Unwin)

To broaden the journal and cope with the increase of submissions, a three co-editor system was introduced in 1998. Wolfgang Karl Härdle (who continued as the EiC), Joe Newton (from mid 1998), and Antony Unwin (from 1999) took the roles of co-editors. Furthermore, twenty additional AEs were newly invited to *Comput Stat* during

this period. The total number of AEs in 2003 increased to 49 from 26 in 1999 (see the right chart in Figure 2). The number of submissions stabilized at about 75 during this period (see the left chart in Figure 2).

2004–2006 (Härdle — Mori — Symanzik)

Yuichi Mori (from 2004) and Jürgen Symanzik (from 2005) participated in the editorial team as successors of Antony Unwin and Joe Newton, respectively. In this period, the regional AE assignment rule started, in which each of the three editors was responsible for submissions from his geographic area: Joe Newton / Jürgen Symanzik for North America, Yuichi Mori for Asia and Oceania except for Mainland China, and Wolfgang Karl Härdle for all other regions. This system was effective to increase the number of submissions from countries from where no manuscripts had been submitted so far, for example, Asian countries such as Japan, Korea, and Taiwan. Each co-editor assigned papers directly to an AE from his geographic region. Therefore, each co-editor invited additional AEs from his geographic region to initiate the regional review process (see the increase of AEs during this time period in the right chart in Figure 2).

Comput Stat also installed its first electronic editorial system in this period, in which editors, AEs, reviewers, and authors communicated with each other for the review-revise processes electronically via a web-based system. From April 2005 to October 2006, *Comput Stat* used “*Editorial Express*” (EE). The introduction of this full electronic editorial system to *Comput Stat* was the result of a long discussion with Springer. In fact, *Comput Stat* was one of the first few journals that switched to such an electronic editorial system.

2007–2011 (Leisch — Mori — Symanzik)

In the middle of 2006, Friedrich Leisch started to take over the EiC role from Wolfgang Karl Härdle and fully replaced him as EiC in 2007. Furthermore, a different electronic editorial system “*Editorial Manager*” (EM) was installed in November 2006. The transition from EE to EM was necessary as Springer started to streamline the submission and publication process for all of its journals in order to speed up the overall publication process. In the remainder of this section, we summarize the trends of submissions, focusing on the number, countries, and topics of the submitted manuscripts, based on data stored in EM.

In this period, the number of submissions increased dramatically: from 75 in 2005 to 220 in 2010 (see the left chart in Figure 2). According to this increase, the number of AEs was increased from 49 in 2003 to 94 at the end of 2011 (see the right chart in Figure 2).

Since the inception of EM in 2006, submissions were received from 66 different countries. A comparison of country of origin of the corresponding author for each paper reveals that in 2007 (first full year of EM usage), submissions were received from 33 different countries, while in 2010 (final complete year of EM usage), submissions were

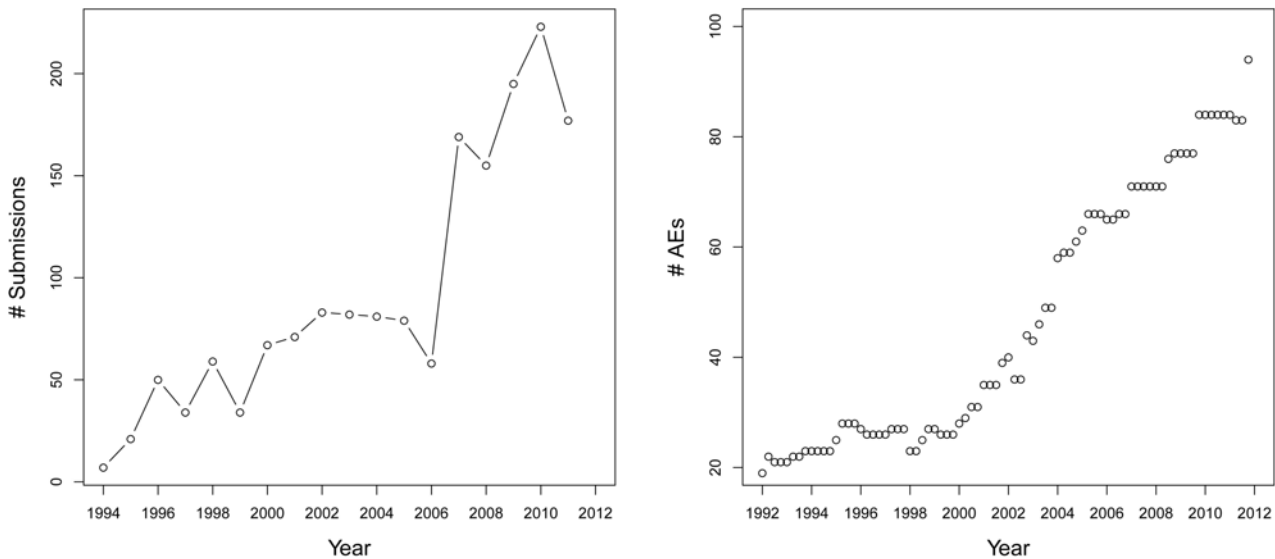


Figure 2: Change in the number of submissions from 1994 until 2011 (left) and change in the number of Associate Editors from 1992 until 2011 (right). The numbers of submissions are based on personal records of the co-editors and electronic records from EE and EM. The numbers of AEs have been derived from the cover pages of the print issues of *Comput Stat*.

received from 41 different countries. In 2007, most submissions were received from corresponding authors residing in the United States (27), India (14), Turkey (12), China (10), and Taiwan (10). In 2010, most submissions were received from Germany (26), United States (23), China (15), Taiwan (14), and Spain (12). Figure 3 shows maps of the submissions by country for 2007 (left) and 2010 (right). Considering this situation, we decided to stop the regional AE assignment rule and we started the pooling of AEs in the spring of 2007. As EM allows each co-editor to access the assignment status of all AEs, AEs could now be assigned to a submission from anywhere in the world that more closely matches their own research interests, and a much better load balance for each AE could be obtained (at least theoretically).

EM allows authors to choose from 509 different classification topics for their manuscripts, ranging from *Adaptive Design* to *WWW (World Wide Web)*. At least one classification must be selected (even though in 2010, three manuscripts made it into the system without a single topic). Occasionally, eight or more classifications were selected, with a record of 16 classifications selected for a single manuscript in 2010.

In 2007, a total of 724 classifications (258 unique) were selected, which corresponds to an average of 4.3 classifications per manuscript. In 2010, a total of 744 classifications (239 unique) were selected, which corresponds to an average of 3.3 classifications per manuscript. It should be noted that similar terms such as *Maximum Likelihood*

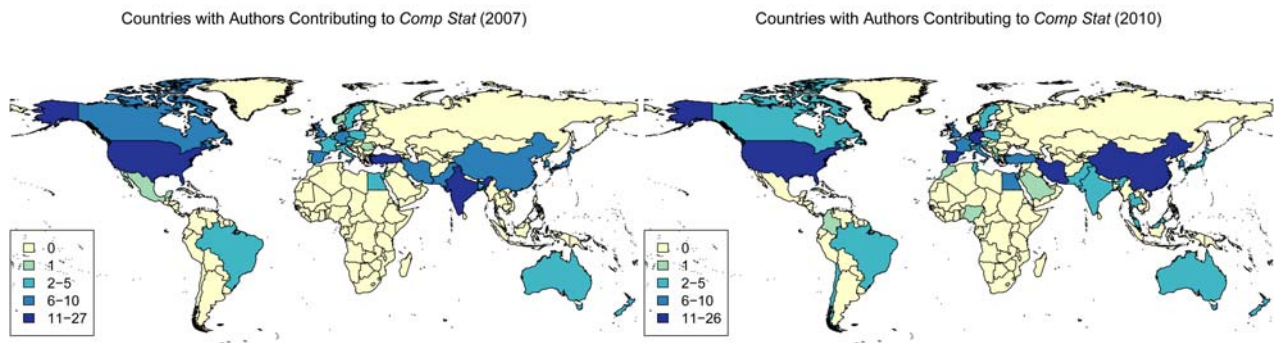


Figure 3: Maps showing the number of submissions per country for 2007 (left) and 2010 (right). It is noteworthy that most additional countries with submissions to *Comput Stat* were located in Asia, Africa, and South America.

and *Maximum Likelihood Estimator*, *Bayesian Analysis* and *Bayesian Estimation*, or *Simulation* and *Simulation Study* often were used as pairs.

Most popular classification topics (used in 9, respectively 10, or more manuscripts) in 2007 and 2010 are listed in Table 1. Little changes with respect to the most popular classification topics occurred. In fact, six of the 10 most popular classification topics used in 2007 are among the 12 most popular classification topics used in 2010. However, topics such as *Mathematical Statistics*, *Inferential Statistics*, and *Reliability* became more popular in 2010, whereas topics such as *Software: R* and *Algorithms for Model Selection* were used for much fewer manuscripts. It remains to be seen in future years whether these numbers indicate any long-term shifts or just are random fluctuations from one year to another.

In October 2011, the new co-editor team, consisting of Yuichi Mori (Editor-in-Chief), Jürgen Symanzik, and Helmut Herwartz, started its work. It should be noted that the current number of 94 AEs (see the right chart in Figure 2 for the overall development) likely will drop again in 2012 as several new AEs already were invited and joined the team of existing AEs during the last few weeks of 2011 while the terms of several current AEs are about to expire at the end of 2011.

SPECIALTIES

Comput Stat has published special articles in irregular time intervals. Noteworthy is our “*Interview with . . .*” series where we have interviewed distinguished researchers in the fields of computational statistics and/or statistical computing. We have published interviews with four researchers so far, selected to maintain some regional balance, that is, Professor Dr. James E. Gentle (George Mason University, USA: Vol. 19, No. 1, 2004), Professor Dr. Antony Unwin (University of Augsburg, Germany: Vol. 20, No. 1, 2005),

Table 1 Number of most popular topics (article classifications) in 2007 and 2010. Only topics with occurrences in 9 (in 2007), respectively 10 (in 2010), or more manuscripts are listed in this table. The numbers in parentheses indicate the ranks in each year for the top-10 topics in each year.

Topic	2007	2010
Monte Carlo Simulation	11 (6)	22 (1)
Statistical Computing	18 (2)	18 (2)
Simulation Study	14 (4)	17 (3)
Data Analysis	20 (1)	16 (4)
Mathematical Statistics	7	15 (5)
Regression Analysis	13 (5)	14 (6)
Inferential Statistics	2	12 (7)
Simulation	8	11 (8)
Bayesian Computing	6	10 (9)
Bayesian Inference	7	10 (9)
Reliability	3	10 (9)
Time Series	11 (6)	10 (9)
Software: R	16 (3)	6
Algorithms for Model Selection	11 (6)	4
Maximum Likelihood Estimator	10 (9)	6
Bayesian Analysis	9 (10)	8

Professor Dr. Genshiro Kitagawa (Institute of Statistical Mathematics, Japan: Vol. 21, No. 1, 2006), and Professor Dr. Andreas Buja (The Wharton School, University of Pennsylvania, USA: Vol. 23, No. 2, 2008). An article with a joint interview of Dr. Robert Gentleman and Dr. Ross Ihaka, the original developers of the R software, will be forthcoming in 2012. In order to make clear what computational statistics is about, we have asked these researchers “Where do you see the roots of Computational Statistics?”, “What was the most prominent development for Computational Statistics in the last five, ten, twenty, thirty years?”, “What do you think of today’s statistical software packages?”, and “Is there a future for Computational Statistics?”, and other, more specialized questions that matched the background of each interviewee.

Special Issues (SIs) are another specialty of *Comput Stat*. We have published articles that were submitted under specific topics as a SI. We invited guest editor(s) for the specific topics and guest editor(s) gathered authors. Also, some SIs were based on sessions from conferences or even on entire conferences or workshops. The reviewed articles were published in a single issue or as a part of a regular issue. All topics for SIs and their guest editors are listed in Table 2. Special issues that are planned for the next two years include *Proceedings of Reisenburg 2010 & 2011*, *Sparse Matrices*, and the *2011 Data Exposition Competition of the American Statistical Association*. While some topics reoccured over time, other topics were unique.

Table 2 List of titles and guest editors of Special Issues.

Year	Vol(No)	Title	Guest Editors
2011	26(4)	<i>Data Viz VI (2008)</i>	A.F.X. Wilhelm, L. Linsen
2011	26(3)	<i>Proceedings of Direction in Statistical Computing 2009</i>	P. Dalgaard, S. Højsgaard
2011	26(2)	<i>Proceedings of Reisenburg Workshop 2009</i>	F. Leisch
2010	25(4)	<i>2006 Data Exposition Competition of the American Statistical Association</i>	P. Murrell
2009	24(2)	<i>Proceedings of Direction in Statistical Computing 2007</i>	P. Murrell
2007	22(4)	<i>Workshop Data and Information Visualisation 2006</i>	S. Klinke
2007	22(3)	<i>Modelling Functional Data in Practice</i>	M.J. Valderrama
2007	22(2)	<i>Partial Least Squares</i>	T. Aluja-Banet, V. Esposito Vinzi, M. Tenenhaus
2006	21(2)	<i>Interval Data</i>	F. Palumbo
2004	19(1)	<i>Papers on Data Visualisation — What can we see in them?</i>	A. Unwin, M. Theus
2003	18(3)	<i>Computational Finance</i>	H. Herwartz
2003	18(2)	<i>Non-parametric modelling (Euroworkshop on Statistical Modelling)</i>	G. Kauermann
2002	17(3)	<i>Statistical Software in the Internet Age</i>	Y. Mori, J. Nakano, Y. Yamamoto
2001	16(3)	<i>Data Mining and Statistics</i>	H. Hofmann, A. Unwin, A. Wilhelm
2000	15(3)	<i>Bayesian Computational Statistics</i>	I.V. Mechelen, P. de Boeck, H. Hoijtink
2000	15(1)	<i>New Techniques and Technologies for Statistics</i>	C. Lauro, V. Esposito Vinzi
1999	14(1)	<i>Interactive Graphical Data Analysis</i>	D.F. Swayne, S. Klinke
1998	13(1)	<i>Strategies for Data Analysis</i>	A. Unwin, M. Theus
1997	12(1)	<i>10 Years AG GLM</i>	G. Osius, G. Tutz
1996	11(4)	<i>Computeraided Analysis of Spatial Data</i>	R. Ostermann
1995	10(1)	<i>Workshop on Modern Methods of Classification</i>	—

THE FUTURE

It is difficult to predict what will happen with *Comput Stat* in the future. Very likely, it will be affected by changes in the overall publication environment. As described in the previous sections, digital, web-based approaches replaced the classical review and publication process. It can be expected that this process will continue even further.

As reported to us by Springer, electronic subscriptions nowadays surpass subscriptions to the print issue of *Comput Stat*. This already benefits many figures that are available in color at no extra cost for the authors in the electronic issue, while color figures in the print issue require a considerable co-payment to be made by the authors. Assuming that this trend is going to continue, one can anticipate that many additional documents will be made available in the online issue in the future, starting with computer code and data (as is already the cases for some other journals), continuing with high-resolution figures and animated movies or with interactive sequences, and possibly up to true 3-dimensional graphics and figures.

The download statistics, shown in Figure 4, were obtained from the website of *Comput Stat* (<http://www.springer.com/statistics/journal/180>) on November 28, 2011. While on peak days more than 150 articles from *Comput Stat* were downloaded from the Springer web site (with a maximum of 410 articles on September 2, 2011 — based on the available 90-day data from August 31, 2011, to November 28, 2011), the number of downloads averaged around 75 articles on regular workdays. Some of the most popular articles were downloaded more than 100 times during this 90-day time period. As more universities and other researchers worldwide will get access to the internet, it can be expected that the overall daily average of downloaded articles will increase in the future, and, therefore, most popular articles will receive even more attention. Perhaps the download information by itself may become more valuable for well-read authors with respect to promotion and tenure decisions than the number of citations that often lag by a few years.

Conclusion

Comput Stat has considerably grown with respect to submissions, number of Associate Editors, and special issues. With a recently updated editorial team, we are looking forward to a promising future and we would like to invite submissions from researchers from around the world for articles related to the numerous aspects of computational statistics.

Finally, we would like to thank our five other co-editors, our more than 130 AEs, our book review and software review editors, all of our editorial assistants and other support team at Springer, the guest editors of our SIs, our thousands of reviewers, our authors, and our readers for their ongoing support of *Comput Stat* since 1992.

Cross-References

Journal of Computational and Graphical Statistics, Computational Statistics and Data Analysis (CSDA), Journal of Statistical Software

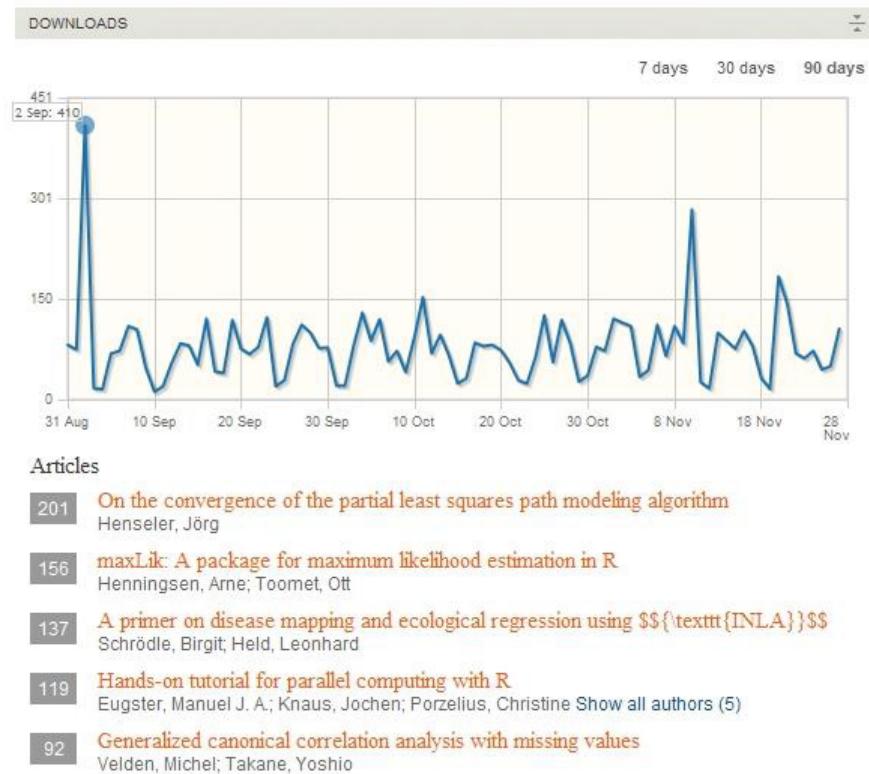


Figure 4: Download statistics from <http://www.springer.com/statistics/journal/180>, obtained on November 28, 2011. Shown are the numbers of downloaded articles for each day over a 90-day period and the five manuscripts with most downloads during this time period.

SFB 649 Discussion Paper Series 2012

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "HMM in dynamic HAC models" by Wolfgang Karl Härdle, Ostap Okhrin and Weining Wang, January 2012.
- 002 "Dynamic Activity Analysis Model Based Win-Win Development Forecasting Under the Environmental Regulation in China" by Shiyi Chen and Wolfgang Karl Härdle, January 2012.
- 003 "A Donsker Theorem for Lévy Measures" by Richard Nickl and Markus Reiß, January 2012.
- 004 "Computational Statistics (Journal)" by Wolfgang Karl Härdle, Yuichi Mori and Jürgen Symanzik, January 2012.

SFB 649, Spandauer Straße 1, D-10178 Berlin
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

