FM 8:
Statistics and Exoplanets

Introduction

Suzanne Aigrain[1] and Eric Feigelson[2]

[1]Department of Physics, University of Oxford, Keble Road, Oxford, OX3 9UU, UK
email: `suzanne.aigrain@astro.ox.ac.uk`

[2]TBD

email: `Department of Astronomy & Astrophysics and Center for Astrostatistics, Penn State University, 525 Davey Laboratory, University Park PA 16802 USA`

Abstract. The IAU's *Statistics and Exoplanets* Focus Meeting brings together observers, modelers and methodologists to discuss the intricate challenges of extracting and interpreting faint planetary signals from dominant starlight. Initiated by the IAU's new groups concentrating on astroinformatics and astrostatistics, the meeting stimulated the wider exoplanetary community as well as experts in data and science analysis. This proceedings presented selected papers from the Focus Meeting.

Keywords. methods: data analysis, methods: statistical, (stars:) planetary systems

1. Motivation

The discovery and characterization of exoplanets requires both superbly accurate instrumentation and sophisticated statistical methods, to extract weak planetary signals from dominant starlight, very large samples and noisy datasets. While numerous exoplanet conferences take place each year, these mostly focus on observational results, their physical implications, and current or future instrument developments; the statistical aspects of the papers presented, while important, are not usually central to the program. We felt it would be timely to hold a meeting dedicated specifically to the statistical work underpinning much of exoplanet science – from the detection of tiny signals buried in correlated noise to the robust inference of planet demographics from diverse, incomplete and often biased surveys. Encouraged by the IAU's Working Group on Astrostatistics and Astroinformatics (now evolved into the Commission B.3 on Astroinformatics and Astrostatistics), we therefore proposed *Statistics and Exoplanets* as a Focus Meeting for the 2015 General Assembly of the IAU. Hot on the heels of the first astro-statistics IAU Symposium – *Statistical Challenges in 21st Century Cosmology*, IAU Symposium #306, Lisbon, 2014 – this would, we hoped, help engage a fruitful discussing between astronomers and statisticians, disseminate and foster best practice, and lead to new collaborations and novel applications of state-of-the-art statistical methods to exoplanetary data and science.

2. Overview

Perhaps the most important intention for this Focus Meeting was to bring astronomers working on exoplanets face to face with mathematical and computational statistics experts. This enables the astronomers to discuss and address key problems they encounter when analysing their datasets, on the one hand, and relevant statistical methods, on the other, with a view to making definite progress on some key challenges that are at the forefront of the field today, such as the robust detection of exoplanets signals in time-series data affected by correlated noise, and the estimation of the incidence of Earth-sized

planets in the habitable zones of Sun-like stars, known as η-Earth. We sought to achieve this by inviting a number of eminent professional statisticians, whom we knew to be interested in engaging with data from a variety of disciplines, but would bring a distinct perspective to the problems regularly faced by exoplanet astronomers.

The speakers at the Focus Meeting oral sessions are listed in the table below with their titles. They include astrostatisticians with an established track record of work on exoplanets, as well as observational astronomers and modelers of planetary systems. They provided overviews of the principal statistical challenges associated with various subfields of exoplanetary research. Together with a selection of contributed talks, the talks were arranged over six sessions covering:

- statistician's perspective
- planet demographics and η-Earth
- planetary signals in sparse datasets
- planetary signals in continuous light curves
- high contrast imaging
- characterisation of exoplanet atmospheres

The final session included a panel discussion with audience participation to reflect on the highlights, lessons learned and ideas expressed during the meeting.

The talks were complemented by a well-populated poster session of $\sim$75 contributed papers. These often presented fascinating results using the latest methodological techniques. The Focus Meeting was capped by a 'Hack Day' splinter session where the authors of software packages designed for the analysis of exoplanet datasets. These informal presentations presented important new methods to potential users, who then had the opportunity to try them out on the spot.

The Focus Meeting was extremely well attended, with from 180 to 130 IAU members at the oral sessions. This is a testament both to the buoyant nature of exoplanets as a field of study and to the timely nature of a meeting concentrating on statistical methods. Methodology developed for one application is often relevant to others, and thus the meeting appealed to a wide range of IAU astronomers. The standard of the talks was uniformly high and the feedback received by the Scientific Organizing Committee very positive. The talks were filmed and archival video posted on the meeting website, `www.exostats.org`, where the interested reader will also find some of the slides of the talks presented during the Focus Meeting and Hack Day.

3. Panel discussion

The meeting concluded with a panel discussion between Wesley Traub (Jet Propulsion Laboratory, US), Thomas Loredo (Cornell University, US), Shay Zucker (Tel Aviv University, IL) and Suzanne Aigrain (Oxford University, UK), chaired by Andrew Collier Cameron (University of St. Andrews, UK). In his opening remarks, the chair highlighted some 'meta-issues' which arose several times during the meeting:

- exoplanet studies often work close to the detection limit, making the need to understand selection functions all the more pressing,
- astronomers seeking sophisticated statistical methods have a tendency to 'reinvent the wheel'. This partly arises from a difficulty with statistical jargon; how do you search for a solution to a problem if you do not know how to describe your problem in the language of those who might be able to solve it?
- reproducibility of reducing and modeling observational data is important. We need systematic and standardised means of providing intermediate results alongside formal publications.

The discussion then focussed on a few topics that were recurrent throughout the meeting such as: (a) the measurement of η-Earth and the associated problem of defining what we mean by an Earth-like planet; and (b) the detection of an unknown number of planetary signals in sparse (radial velocity) datasets 'decorated' with correlated noise. All panelists recognised that the tremendous progress in the last few years depends critically on the use of powerful statistical procedures that are becoming standard practice among exoplanetary researchers. But the panelists also agreed that we have much more to learn from the statistical community, as well as from the literature in other fields ranging from electrical engineering to econometrics.

4. Conclusion and thanks

We feel this was a very successful and enjoyable Focus Meeting and hope that it will be the first in continuing cross-disciplinary interchanges statistical aspects of exoplanet studies. The management and archival of the meeting website `www.exostats.org` is being taken over by Penn State University, and the intention is that it should be used by future meetings on similar topics too.

The Scientific Organizing Committee was excellent in structuring the meeting and selecting speakers. Committee members were Suzanne Aigrain (University of Oxford, UK, co-chair), Andrew Collier-Cameron (University of St Andrews, UK), Laurent Eyer (Observatoire de Geneve, CH), Eric Feigelson (Penn State University, US, co-chair), Philip Gregory (University of British Columbia, CA), Chris Koen (University of Western Cape, ZA), Michael Liu (University of Hawaii, US), Oleg Malkov (Russian Academy of Science, RU), Claire Moutou (University of Marseille-Provence and CFHT, FR), Sascha Quanz (ETH Zurich, CH), and James Berger (Duke University, US).

We thank the IAU Working Group on Astrostatistics and Astroinformatics, under the leadership of Eric Feigelson and Prajval Shastri (Indian Institute for Astrophysics, IN), for stimulating the organization of this meeting. Joseph Hilbe (Arizona State University US) represented the International Astrostatistics Association affiliated with the International Statistical Institute (sister organization to the IAU) that cosponsored the meeting. James Berger, G. Jogesh Babu (Penn State University, US), Jessi Cisewski (Yale University, US) and Joseph Hilbe are professors of statistics who attended and stimulated the meeting. Andrew Collier-Cameron moderated the panel discussion during the final session. Andrew Howard (University of Hawaii US) and Michael Liu assisted as the Local Organizing Committee. Paul Wilson (Institut d'Astrophysique FR) developed the superb exostats.org Web site and arranged the recording, broadcasting, and archiving of oral talks.

FM8 *Statistics and Exoplanets* Speakers

Title	Presenter
Transit Timing Variations as a Tool for the Bayesian Characterization of Exoplanets	Eric B. Ford
Overview of modern Bayesian statistical methods	James Berger
Planet Demographics from Transits	Andrew Howard
Kepler Reliability Metrics and Their Use in Occurrence Rate Calculations	Steve Bryson
The Occurrence of Earth-Like Planets Around Other Stars	Will M. Farr
Hierarchical inference for exoplanet populations	Daniel Foreman-Mackey
Bayesian planet searches in radial velocity data	Phil Gregory
Wide Giant Planets are Rare: Planet Demographics from Direct Imaging	Beth Biller
The Various Challenges of Subtracting Speckles and Planet Detection/Characterization in High Contrast Imaging	Christian Marois
Planet Frequency beyond the Snow Line from MOA-II Microlensing Survey	Daisuke Suzuki
Astrometric exoplanet surveys in practice: challenges, opportunities, and results	Johannes Sahlmann
Dealing with activity in RV planet searchers	Isabelle Boisse
Estimations of uncertainties of frequencies	Laurent Eyer
Significance of noisy signals in periodograms	Maria Süveges
Advances in the Kepler Transit Search Engine and Automated Approaches to Identifying Likely Planet Candidates in Transit Surveys	Jon M. Jenkins
Combining Transit and Radial Velocity Data to Infer the Planet Mass-Radius-Flux Distribution	Leslie A. Rogers
BART: A Probabilistic and automated tool for the vetting of transits	Olivier Demangeon
Validation of transting planet candidates: a Bayesian view	Ridrigo F. Diaz
Probabilistic Mass-Radius Relationship for Sub-Neptune-Sized Planets	Angie Wolfgang
Probabilistic stellar rotation periods with Gaussian processes	Ruth Angus
A population-based Habitable Zone perspective	Andras Zsom
Reliable extraction of transmission and emission spectra using deterministic and stochastic systematics models	Neale Gibson
Overcoming Degeneracies in Exoplanet Spectra	Björn Benneke
Measuring Transmission Spectra from the Ground	Andres Jordan
Approximate Bayesian Computation	Jessi Cisewski
Analyzing Complex and Structured Data via Unsupervised Learning Techniques	Kai L. Polsterer