

Overview of the WGAA's
State-of-the-Profession White Paper
submitted to Astro2020

arxiv.org/pdf/1909.11714.pdf



UNIVERSITY OF
TORONTO

Gwendolyn Eadie
Assistant Professor of Astrostatistics

Overview of the WGAA's State-of-the-Profession White Paper submitted to Astro2020

arxiv.org/pdf/1909.11714.pdf

PI: Gwendolyn Eadie

Co-authors:

Thomas Loredo, Ashish A. Mahabal, Aneta Siemiginowska, Eric Feigelson, Eric B. Ford, S.G. Djorgovski, Matthew Graham,, Zeljko Ivezic, Kirk Borne, Jessi Cisewski-Kehe, J. E. G. Peek, Chad Schafer, Padma A. Yanamandra-Fisher, C. Alex Young

Key Points

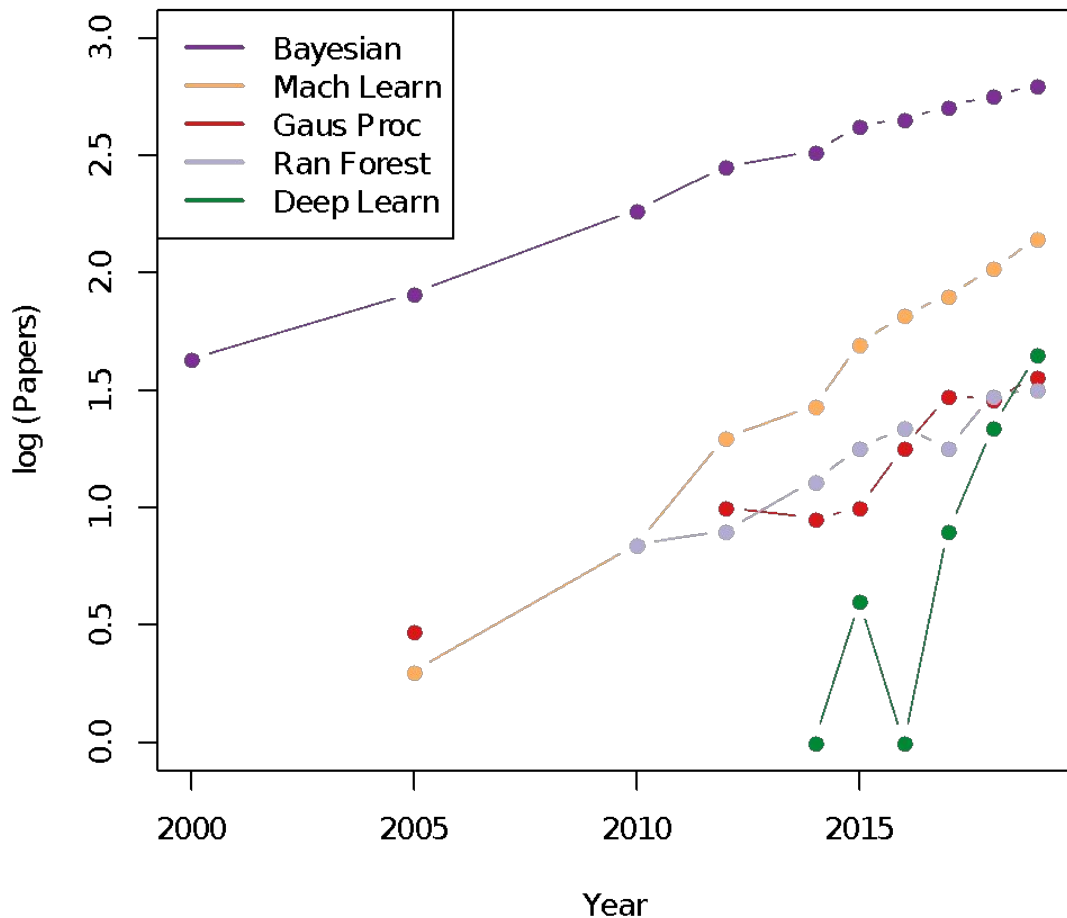
In our State-of-the-Profession
White Paper

- **Highlights strong and growing impact** of Astrostatistics & Astroinformatics
- **Identifies issues** hampering the growth of this new field
- **offers recommendations** for improved support of both research and education in Astrostatistics & Astroinformatics

Advances in statistics and machine learning methods making an impact in astronomy

- Hierarchical Bayesian analysis and Bayesian computation
- Nonlinear dimensionality reduction
- Deep learning

Papers in AAS Journals



\log_{10} (papers) per year using statistical and machine learning methods since 2000.

(Figure 1 in white paper)

Advances in statistics and machine learning methods are making an impact in astronomy

- Hierarchical Bayesian analysis and Bayesian computation
- Nonlinear dimensionality reduction
- Deep learning

Many of these advances have occurred rapidly, creating **unfamiliar challenges and imbalances for the educational, funding, and quality control structures** of the field

Key Issues: Education, funding, and quality control

1. Education Gap
2. Funding Gap
3. Quality Control Gap

The Education Gap

- Astronomers typically well-trained in mathematics ...
 - but typically not well-trained in applied mathematics, statistics, and computer science relating to extraction of reliable information from complex, noisy datasets

The Education Gap

- Astronomers typically well-versed in programming ...
 - but not ready for the challenges of Big Data, including data storage, access, and efficient on HPC

(see also Besla et al 2019 white paper “Astro2020: Training the Future Generation of Computational Researchers”

<https://arxiv.org/abs/1907.04460>

The Funding Gap

- Grants to improve methodologies are scarce in astronomy ...
 - but other sciences disciplines do receive funding for such work! (e.g. biostatistics, statistics and informatics for Earth sciences)

The Quality Gap

- High standards for analysis methodology are not consistently set for e.g., publications ...
 - at the same time there is increased demand for people who can peer-review such papers

Recommendations

Strategic Plan / Recommendations

Astro 2020

Decadal Survey on Astronomy and Astrophysics

The National
Academies of
SCIENCES
ENGINEERING
MEDICINE



- Universities
 - Revise the undergraduate and graduate curriculum
 - Financially support interdisciplinary summer schools
 - Financially support interdisciplinary PhD programs
 - Create interdisciplinary postdoctoral appointments
 - Establish permanent appointments for data science in astronomy
 - Establish permanent joint-appointments with relevant departments (e.g. statistics, computer science, information sciences, etc.)

Strategic Plan / Recommendations

Astro 2020

Decadal Survey on Astronomy and Astrophysics

The National
Academies of
SCIENCES
ENGINEERING
MEDICINE



- NSF and NASA
 - Provide grant support for research in astrostatistics and astroinformatics
 - Encourage reviewers of postdoctoral applications to pay special attention to astrostatistics & astroinformatics researchers
 - Financially support research regarding data infrastructure
 - Instate a 3-year interdisciplinary fellowship program
 - Support interdisciplinary research collaborations that seek to include substantive contributions from experts outside astronomy

Astro 2020

Decadal Survey on Astronomy and Astrophysics

*The National
Academies of* | SCIENCES
ENGINEERING
MEDICINE



Astro 2020

Decadal Survey on Astronomy and Astrophysics

*The National
Academies of* | SCIENCES
ENGINEERING
MEDICINE



Canada has an astronomy decadal review too!

Astro 2020

Decadal Survey on Astronomy and Astrophysics

*The National
Academies of* | SCIENCES
ENGINEERING
MEDICINE



Canada has an astronomy decadal review too!

Long Range Plan 2020 (LRP2020)

LRP2020 White paper: *Astrostatistics in Canada*

PI: Eadie, G.

Co-authors (astronomers *and* statisticians):

Arash Bahramian, Pauline Barmby, Radu Craiu, Derek Bingham, Renée Hložek, JJ Kavelaars, David Stenning, Samantha Benincasa, Guillaume Thomas, Karun Thanjavur, Jo Bovy, Jan Cami, Ray Carlberg, Sam Lawler, Adrian Liu, Henry Ngo, Mubdi Rahman, Michael Rupen

LRP2020 White paper: *Astrostatistics in Canada*

Eadie, G. + 18 co-authors 2019 (astronomers and statisticians)

Table 1: Astrostatistics and Astroinformatics Centres, Working Groups, Associations, etc.

Association/Group	Formed	Activities and/or Objective
Centre for Astrostatistics	2003	organise an annual Summer School in Statistics for Astronomers, astrostatistics research, hosts the Astrostatistics & Astroinformatics Portal https://asaip.psu.edu/
LSST's Informatics & Statistics Science Collaboration	2009	develop tools for large astronomical surveys
AAS WGAA	2012	organise sessions and panels at AAS conferences; advocate for curricula change
IAA	2012	foster collaborations between statisticians and astronomers
Astrostatistics Facebook Group	2013	over 4300 members, active discussion
ASA AIG	2014	increase statisticians' involvement in astronomy research; propose and organise sessions at the annual Joint Statistical Meetings; annual student paper competition
IEEE Astrominer Task Force	2014	contribution to machine learning, data-mining, and computational methods
IAU Commission B3	2015	Commission on Astroinformatics and Astrostatistics (previously Working Group since 2012)
IAIA	2019	host the annual astroinformatics conference

Survey results

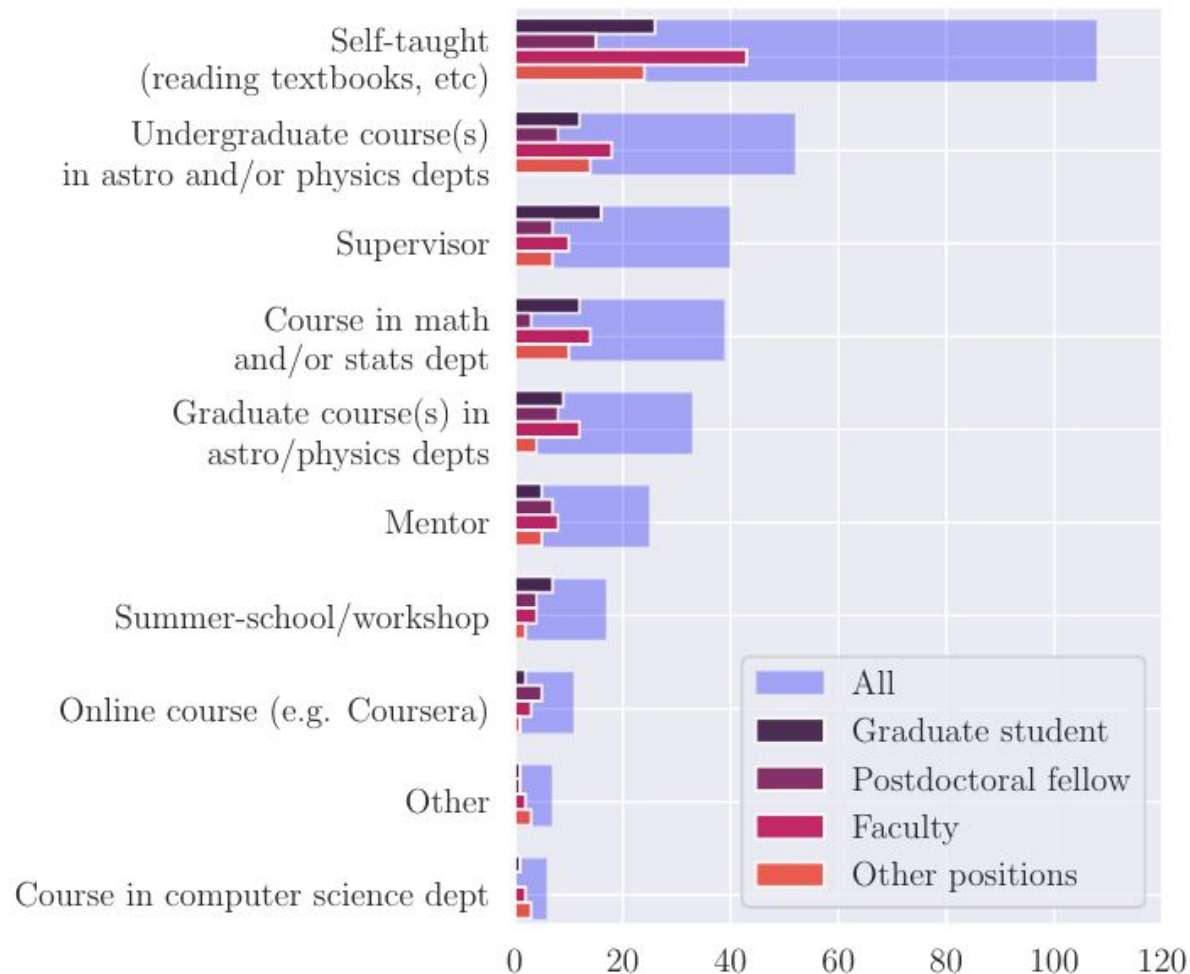
Eadie, G. + 18 co-authors (2019)

*“When/From whom did
you receive training in
statistics?”*

Survey results

Eadie, G. + 18 co-authors (2019)

“When/From whom did you receive training in statistics?”



Survey results

Eadie, G. + 18 co-authors (2019)

“Do you wish you had received received more training in statistics at the undergraduate, graduate, and postdoc levels?”

Survey results

Eadie, G. + 18 co-authors (2019)

“Do you wish you had received received more training in statistics at the undergraduate, graduate, and postdoc levels?”

“YES” responses:

85%	<i>wished they had received training in undergrad</i>
82%	<i>“ “ “ “ “ “ in grad school</i>
36%	<i>“ “ “ “ “ “ during postdoc</i>

Survey results

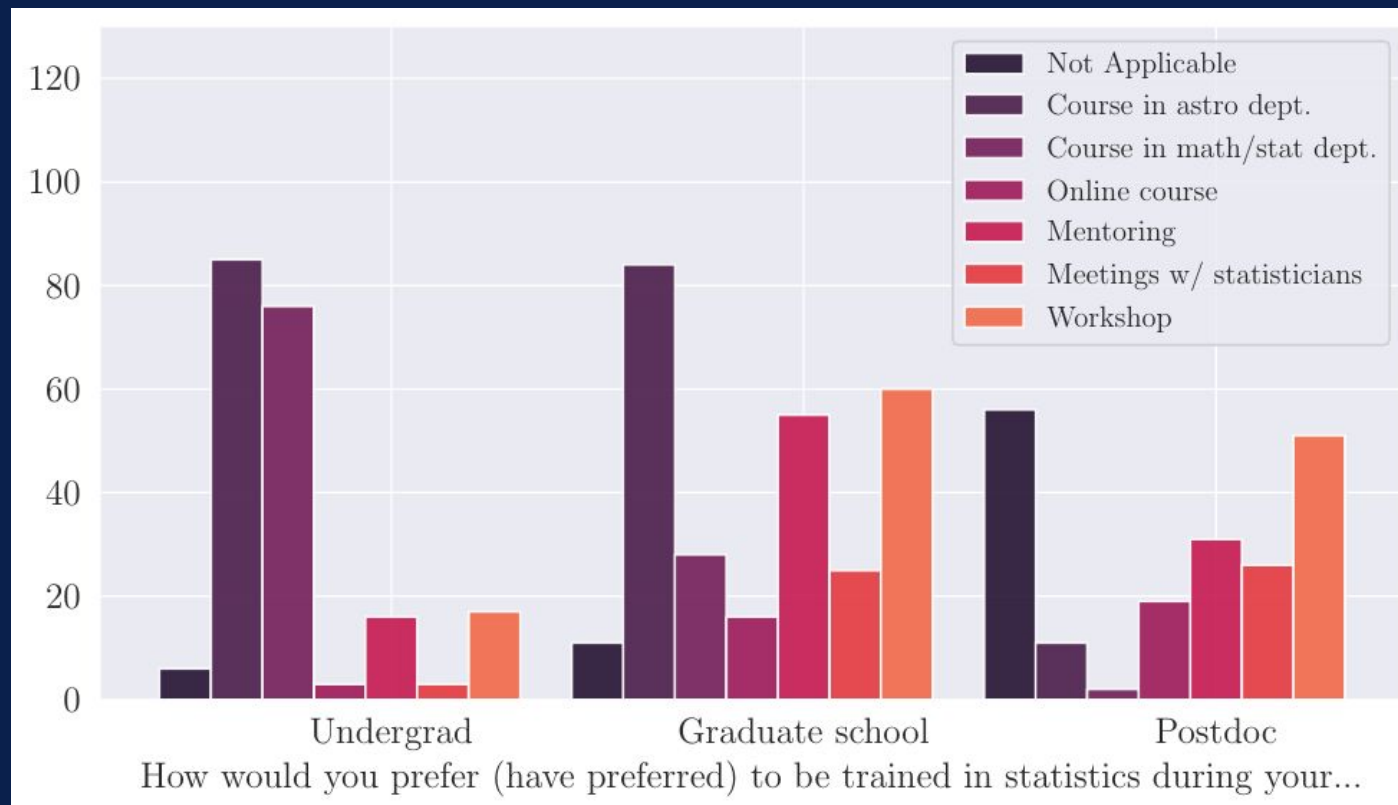
Eadie, G. + 18 co-authors (2019)

“How would you have preferred to be trained in statistics at the undergraduate, graduate, and postdoc levels?”

Survey results

Eadie, G. + 18 co-authors (2019)

“How would you have preferred to be trained in statistics at the undergraduate, graduate, and postdoc levels?”



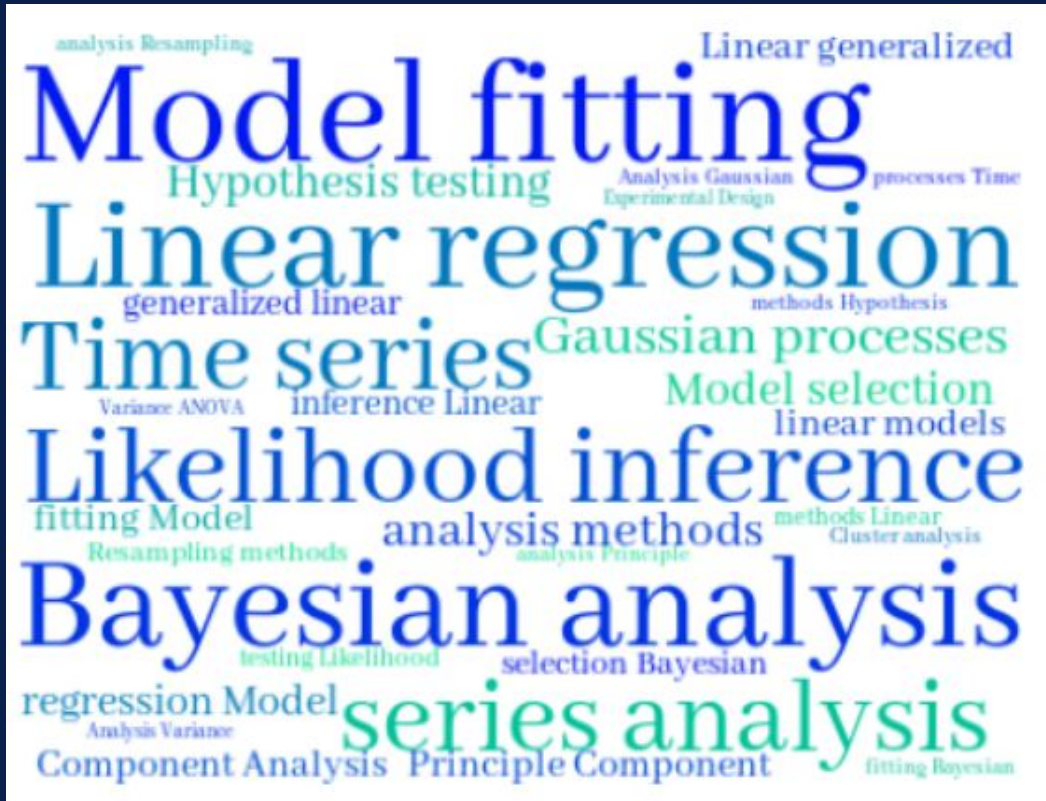
Survey results

Eadie, G. + 18 co-authors (2019)

What are the statistical tools/models you find useful in your work?

Survey results

Eadie, G. + 18 co-authors (2019)



What are the statistical tools/models you find useful in your work?

Survey results

Eadie, G. + 18 co-authors (2019)

*What statistical
tools/models do you wish
you knew more about?*

What are the main challenges you have encountered when working with statisticians, applied mathematicians, or computer scientists?

Free-form Responses

“

I had zero exposure to stats in physics and astrophysics. I took one undergraduate course in an engineering department, but most of my stats skills are learned in industry when needed. This has been my biggest challenge in industry. Stats are extremely important in analyses and making informed business decisions and should be emphasized much, much more in physics departments!

”

Free-form Responses

“

*I would be a much stronger researcher, and be able to produce much more meaningful and exciting science, if I had received proper training in statistics. Instead, I had a series of upper undergrad and graduate courses that *assumed* we knew the basics of stats, but never actually taught us any of it. I've only had one formal data analysis course, and because our education had been so lacking up to that point, the course tried to fit too much into too little time. I still would not be confident doing any serious statistical analysis.*

”

Recommendations for the LRP2020

- Institutional Level
 - e.g. create venues to interact with researchers in methods departments, start interdisciplinary research teams, change curricula for astronomy programs
- National Level
 - Work with the **Statistical Society of Canada (SSC)**
 - Take advantage of opportunities provided by the **Canadian Statistical Sciences Institute (CANSSI)**
- International Level
 - Collaborate with other organizations (e.g. WGAA of the ASA)



Canadian Statistical Sciences Institute Institut canadien des sciences statistiques

*Data • Discoveries • Decisions
Données • Découvertes • Décisions*

“Building on the international stature of the Canadian statistical community, **CANSSI seeks to develop all areas of the statistical sciences, including interdisciplinary research** where statistical innovation is essential to the development of other disciplines.”

- **Kick Start Research Program**
- **Collaborative Research Teams (CRTs), ~180K CAD over 3 years**
- **CANSSI Postdoctoral Fellowships**
www.canssi.ca