

# EMI TANAKA

## WORK EXPERIENCE

- **Lecturer**  
Monash University,  
Department of Econometrics and Business Statistics  
2020/01 - current
- **Lecturer**  
The University of Sydney,  
School of Mathematics and Statistics  
2017/01 - 2019/12
- **Research fellow**  
University of Wollongong,  
School of Mathematics and Applied Statistics  
2014/01 - 2017/01

## EDUCATION

### **PhD, Statistics**

Statistical Methods for Improving Motif Evaluation

Supervisor: Dr. Uri Keich

School of Mathematics and Statistics

The University of Sydney, Sydney, Australia, 2015

### **Bachelor of Science (Advanced Mathematics), Honours I**

Major in Mathematics and Statistics

The University of Sydney, Sydney, Australia, 2010

## PUBLICATIONS

1. Morota, G, Cheng, H, Cook, D, & Tanaka, E (2021) ASAS-NANP SYMPOSIUM: Prospects for interactive and dynamic graphics in the era of data-rich animal science. *Journal of animal science*, 99(2). *Citations: 1.*
2. Tanaka, E (2020) Simple outlier detection for a multi-environmental field trial. *Biometrics*, 76(4), 1374-1382. *Citations: 4.*
3. Tanaka, E & Hui, F (2019) Symbolic formulae for linear mixed models. *Statistics and data science*, 3-21
4. Hui, F, Tanaka, E, & Warton, D (2018) Order selection and sparsity in latent variable models via the ordered factor LASSO. *Biometrics*, 74(4), 1311-1319. *Citations: 8.*
5. Norman, A, Taylor, J, Tanaka, E, Telfer, P, Edwards, J, Martinant, J, & Kuchel, H (2017) Increased genomic prediction accuracy in

## Contact

✉ [emi.tanaka@monash.edu](mailto:emi.tanaka@monash.edu)

🐙 [github.com/emitanaka](https://github.com/emitanaka)

🐦 [@statsgen](https://twitter.com/statsgen)

For more information, please  
contact me via email.

## Skills

**Expert:** R, HTML/CSS, LaTeX

**Intermediate:** Git/GitHub, Python,  
Bash, JS

**Languages:** English (fluent) and  
Japanese (conversational)

## Interests

**Research:** experimental design,  
mixed models, data visualisation,  
bioinformatics, statistical genetics,  
selective breeding

**Non-Research:** drawing (but not  
good at it), reading (manga,  
manhwa and non-fiction books)

*This resume was made with the R package  
**pagedown**.  
Last updated on 2021-06-05.*

wheat breeding using a large australian panel. *Theoretical and applied genetics*, 130(12), 2543-2555. *Citations*: 17.

6. Tanaka, E, Ral, J, Li, S, Gaire, R, Cavanagh, C, Cullis, B, & Whan, A (2017) Increased accuracy of starch granule type quantification using mixture distributions. *Plant methods*, 13, 107. *Citations*: 5.
7. Tanaka, E (2014) Statistical methods for improving motif evaluation. *PhD Thesis*
8. Tanaka, E, Bailey, T, & Keich, U (2014) Improving MEME via a two-tiered significance analysis. *Bioinformatics*, 30(14), 1965-1973. *Citations*: 17.
9. Liachko, I, Tanaka, E, Cox, K, Chung, S, Yang, L, Seher, A, Hallas, L, Cha, E, Kang, G, Pace, H, Barrow, J, Inada, M, Tye, B, & Keich, U (2011) Novel features of ARS selection in budding yeast *lachancea kluyveri*. *BMC genomics*, 12, 633. *Citations*: 22.
10. Tanaka, E, Bailey, T, Grant, C, Noble, W, & Keich, U (2011) Improved similarity scores for comparing motifs. *Bioinformatics*, 27(12), 1603-1609. *Citations*: 50.

Citation counts are sourced from Google Scholar at 2021-06-05.

## SOFTWARE

### edibble

An R-package that implements the grammar of experimental design.

*Creator and maintainer. GitHub stars: 123*

 Source: <https://github.com/emitanaka/edibble>

 Documentation: <https://edibble.emitanaka.org>

### anicon

An R-package to insert animated icons for R markdown and Shiny apps.

*Creator. GitHub stars: 113*

 Source: <https://github.com/emitanaka/anicon>

### nestr

An R-package to build hierarchical or nested structures.

*Creator and maintainer. GitHub stars: 12*


 Source: <https://github.com/emitanaka/nestr>

 Documentation: <https://nestr.emitanaka.org>

### monash

A utility R-package with consolidated tools and templates for staffs at Monash University.

*Creator and maintainer. GitHub stars: 3*

 Source: <https://github.com/numbats/monash>

### xaringan

An R package for creating slideshows with remark.js through R Markdown.

*Contributor. My contribution is the adaptation of the ninja-theme shown in the Documentation. GitHub stars: 1191*

🔗 Source: <https://github.com/yihui/xaringan>

🔗 Documentation: <https://github.com/emitanaka/ninja-theme>

## 📊 gghdr

Plots of highest density regions (HDR) for ggplot2.

*Listed as author but my contribution is small. GitHub stars: 34*

🔗 Source: <https://github.com/ropenscilabs/gghdr>

🔗 Documentation: <https://ropenscilabs.github.io/gghdr>

## 📊 shinycustomloader

Add a custom loader for R shiny.

*Creator. GitHub stars: 101*

🔗 Source: <https://github.com/emitanaka/shinycustomloader>

## 📊 datalegreyar

Datalegreya, the typeface that melts text and data visualisation, for R markdown.

*Creator. GitHub stars: 39*

🔗 Source: <https://github.com/emitanaka/datalegreyar>

## 🔗 Tomtom

A motif comparison tool part of the MEME Suite.

*The -incomplete-scores is an implementation as a result of Tanaka et al. (2011).*

🔗 Documentation: <https://meme-suite.org/meme/doc/tomtom.html>

## 🗣️ TALKS

List of talks (and links to the slides if available) are at <https://emitanaka.org/talks.html>.

Below show the last 10 talks.

### • Visual Inference

FukuokaR

2021-03-27

### • The Grammar of Experimental Design

TokyoR

2021-01-23

### • Advent of “Grammar”: Bridging Statistics and Data Science for the Design of Experiments

Statistical Society of Australia

2020-11-27

### • Rethinking the framework to specify the design of experiments

Department of Agriculture and Fisheries, Queensland

2020-11-11

- **What every young statistician should know:  
Developing your identity, networking and marketing**  
Statistical Society of Australia Western Australian Branch  
2020-09-30
- **Advent of “Grammar”: Bridging Statistics and Data  
Science for the Design of Experiments**  
Monash Bioinformatics Seminar  
2020-07-15
- **Beyond Beamer: Modern and Dynamic Presentations  
with R Markdown**  
Statistical Society of Australia Canberra & NSW Branch  
2020-06-23
- **Software design, selection and estimation for latent  
variable models**  
WOMBAT  
2019-11-28
- **Symbolic model formulae for linear mixed models  
illustrated with the analysis of agricultural data**  
Department of Mathematics and Statistics, Macquarie University  
2019-10-15
- **Symbolic model formulae for linear mixed models  
illustrated with the analysis of agricultural data**  
Research School of Finance, Actuarial Studies and Statistics,  
Australian National University  
2019-09-16







## WORKSHOPS

Workshop	Date	Location	Host
<a href="#">Data Visualisation with R</a>	15-16 Apr 2021	Online	Statistical Society of Australia, Canberra Branch
<a href="#">Data Wrangling with R</a>	1-2 Dec 2020	Online	Statistical Society of Australia, NSW Branch
<a href="#">Data Visualisation with R</a>	11-12 Nov 2020	Online	Statistical Society of Australia, Victoria Branch
<a href="#">Data Visualisation with R</a>	28-29 Jul 2020	Online	Statistical Society of Australia, Victoria Branch
<a href="#">Tidyverse &amp; R Markdown</a>	1 Dec 2019	Adelaide	International Biometrics Society, Australasia Region
<a href="#">Building R Packages &amp; R Markdown</a>	19 Nov 2019	Melbourne	Statistical Society of Australia, Victoria Branch
<a href="#">Communicating with Data via R Markdown</a>	4 Oct 2019	Sydney	COMBINE (subcommittee of ABACBS)
<a href="#">Gaining skills in biostatistical consultancy</a>	4 Jul 2019	Melbourne	Statistical Society of Australia, Biostatistics Section
<a href="#">Statistical Methods for Omics Assisted Breeding</a>	12-15 Nov 2018	Tokyo	University of Tokyo

## TEACHING

Course ID	Course	University	Semester	Median student evaluation score
ETC5521	Exploratory Data Analysis	Monash	2020 S2	4.82 / 5
ETC5523	Communicating with Data	Monash	2020 S2	4.88 / 5
ETC5512	Wild-Caught Data	Monash	2020 S1	4.61 / 5
STAT3012	Applied Linear Models	Sydney	2019 S1	—
STAT3012	Applied Linear Models	Sydney	2018 S1	—
STAT5002	Introduction to Statistics	Sydney	2018 S1	—
STAT3012	Applied Linear Models	Sydney	2017 S1	—
STAT5002	Introduction to Statistics	Sydney	2017 S1	—
STAT906	Experimental Design	Wollongong	2016 S2	—

## SERVICE

-  **President**  
 Statistical Society of Australia, Victoria Branch  
 2021/03-ongoing
-  **Vice President**  
 Statistical Society of Australia, Victoria Branch  
 2023/03-2024/03
-  **Assistant Secretary**  
 Statistical Society of Australia, NSW Branch  
 2017/03-2018/03
-  **Secretary**  
 Statistical Society of Australia, NSW Branch  
 2018/03-2020/03
-  **Social Media Coordinator**  
 International Biometrics Society, Australasian Region  
 2018/01-ongoing
-  **Member**  
 useR! 2021, Program Committee  
 2021/03-2021/03

## PROFESSIONAL MEMBERSHIPS

- Statistical Society of Australia

- International Biometrics Society

## AWARDS & DISTINCTIONS

- **ARC Centre of Excellence**

*Associate Investigator*

ARC Centre of Excellence for Mathematical and Statistical Frontiers (ACEMS)

2020

- **ARC Industrial Transformation Training Centre**

*Chief Investigator*

Data Analytics for Resources and Environments (DARE)

2019

\$3,973,202

- **R Consortium Grant**

*Lead Investigator*

Symbolic Formulae for Linear Mixed Models

2019

\$6,000

- **Sydney Institute of Agriculture Research Project Grant**

*Chief Investigator*

Land surface models of carbon and water do not work in agricultural landscapes where it actually matters they work

2018

\$90,000

- **Faculty of Science Research Equipment and Infrastructure Scheme**

*Chief Investigator*

Sydney bioinformatics and biometrics big data research server

2018

\$36,040