Mitchell O'Hara-Wild

Data Scientist

Melbourne, Australia

mail@mitchelloharawild.com | mitchelloharawild | mitchelloharawild

Current work

Monash University

Clayton, Australia

PhD Candidate

Feb. 2023 - Present

- Research in statistical software for time series analysis and probabilistic forecasting at scale.
- Key research topics include graph-based forecast reconciliation, coherent pruning of aggregation structures, and time-series visualisation
- Underpinning this research are the R packages distributional, graphvec, and mixtime which provide vector-based data structures for distributions, graphs, and time.

Nectric

Melbourne, Australia

Founder & Data Scientist

Jan. 2019 - Present

- Data science consulting projects including spatio-temporal modelling, and large scale forecasting.
- Development of interactive web applications using Shiny for R.
- Hosting workshops improving data literacy, development, analysis and modelling capabilities.

Australia-Aotearoa Consortium for Epidemic Forecasting & Analytics (ACEFA)

Australia

Statistician

Aug. 2020 - Present

- ACEFA provides situational assessment for public health officials in Australia and New Zealand about epidemic diseases.
- I lead the evaluation and ensembling of component forecasts of Influenza, RSV, and COVID-19.
- Recently I have been developing new interactive platforms for stakeholders to engage with and better understand the forecasts.

R Journal

Online

Technical Editor

Jan. 2024 - Present

- The publishing and deployment of articles and issues for the R Journal.
- Maintenance and development of R Journal software, including rjtools, rjdistill, and rj.
- Created software for managing submitted articles and publishing them in both interactive HTML and static PDF.

Monash University

Clayton, Australia

Research Assistant

Oct. 2015 - Present

- Primarily working as a research software engineer creating R packages for time series forecasting techniques.
- Other work involves consulting on large scale time series forecasting and cross-sectional modelling.
- Consulting project clients include: Doherty Institute, Huawei, Monash University, R Consortium, and the NSW Chief Scientist office.

Monash University

Clayton, Australia

Teaching Associate

Mar. 2016 - Present

- Sessional teaching of university students undertaking Bachelor's and Master's degrees.

Qualifications

Monash University

Clayton, Australia

BCom (Hons) in Econometrics

Mar. 2017 - Nov. 2017

- Recipient of the Econometrics Honours Memorial Scholarship, Dean's Honour, Dean's Commendation, and best in class for 5 units.
- Honours research project was to develop a state space model for quickly forecasting time series with multiple seasonalities.
- Studied units include Bayesian and frequentist econometrics, advanced statistical modelling and computational science.

Monash University

Clayton, Australia

BCom & BSc (majoring in Econometrics, Mathematical Statistics, and Computational Science)

Mar. 2013 - Nov. 2016

- Recipient of the Monash Community Leaders Scholarship, International Institute of Forecasters Award, and best in class for 4 units.
- Mentor for the Access Monash Ambassador Program (2015 and 2016)
- Participant of the Vice-Chancellor's Ancora Imparo Student Leadership Program (2014)
- Each of the three disciplines I majored in provided different perspectives for working with data.

Academic contributions

Current Research Areas

The primary research area of my PhD relates to tools and techniques for large-scale time series analysis and forecasting. Large collections of time series are usually structurally related, whereby series of interest are constructed by aggregating other time series. Independent forecasts of these series are adjusted using forecast reconciliation to maintain the aggregation structure. My research extends time series aggregation and forecast reconciliation methods to support arbitrary aggregations using graphs. This work enables non-linear aggregation constraints, cross-temporal reconciliation of sub-daily data, and pruning noisy time series from large coherent time series.

A substantial element in this research is the design of statistical software. Large-scale forecasting involves many different data types beyond time itself. My recent research has focused on developing vectorised data structures for these data types, including distributional (for forecast distributions), mixtime (for temporal aggregation), and graphvec (for representing aggregation relationships). To improve exploratory time series analysis tools, I have also been working on the grammar of temporal graphics and ggtime, which leverages mixtime for better time-series visualisation.

Unrelated to my PhD, I have also been developing scalable teaching tools for self-paced e-learning. This includes moodlequiz for creating Moodle quizzes with literature programming, and online learning modules with programming exercises using webr and qlcheckr. I have also recently been experimenting with using LLMs for exploring and cleaning unstructured data.

Software

My primary research output is the translation of academic research into open source software packages. The software primarily relates to time series forecasting, literate programming, and data visualisation. Some more notable software packages are listed below.

2025 **ggtime**: Lead developer
Extends ggplot2 with a grammer of temporal graphics.

2025 **mixtime**: Lead developer
Vectorised and extensible time classes.

2025 **graphvec**: Lead developer
Vectorised graph nodes and edges.

2025 **vecvec**: Lead developer
Vectors of vctrs for efficient storage of mixed vector types.

2025 **fromhere**: Lead developer
Relative file paths from different places, a declarative version of the here package.

2022 **rjtools**: Lead developer
Tools and templates for publishing articles to the R Journal.

2022 **gghdr**: Author
Extends ggplot2 with plot types for visualising highest density regions.

2020 **distributional**: Lead developer
Vectorised distribution objects with methods for manipulating and computing on probability distributions.

2019 **fable**: Lead developer
A collection of time series models for use in a tidyverse workflow.

2019 **feasts**: Lead developer
A collection of features, decompositions, statistics and graphics tools for the analysing tidy time series data.

2019 **fabletools**: Lead developer
Provides common utilities for the fable forecasting framework, allows users to extend fable with new models.

2019 **tsibbledata**: Lead developer
Diverse datasets in a tidy time series data structure (tsibble). Useful in examples of tidy time series analysis.

2018 **tsibble**: Contributor
Tidy temporal data structures and tools. The data for the tidy time series collection of packages.

2018 **vitae**: Lead developer
Dynamically generate a Résumé or CV using R Markdown. This CV is created using code with that package!

2018 **taipan**: Lead developer
Generates shiny apps for annotating image data, which is useful for training machine learning models.

2017 **fasster**: Lead developer
Implementation of the FASSTER model for forecasting complex multiple seasonal patterns.

2017 **icons**: Lead developer
Embed SVG icons in R documents such as slides, reports and apps.

2017 **ggquiver**: Lead developer
Extends ggplot2 for displaying vector fields on plots.

2015 **forecast**: Author
Methods and tools analysing univariate time series data and producing model-based forecasts.

Presentations

2025 **Probabilistic forecast combinations**: Introduction to ensemble, decomposition, and coherent forecasting
`https://slides.mitchelloharawild.com/acefa-launch/`

2025 **Designing ggtime: A grammar of temporal graphics**: A discussion about ggplot2 extension design.
`https://slides.mitchelloharawild.com/ggextenders-ggtime/`

2024 **Scalable self-paced e-learning with automated feedback**: Using e-learning for teaching at scale.
`https://slides.mitchelloharawild.com/wombat2024/`

2024 **Statistical computing with vectorised operations on distributions**: Introducing distributional for R.
`https://slides.mitchelloharawild.com/user2024/`

2024 **Feature based graph pruning for improved forecast reconciliation**: Coherent large scale forecasting.
`https://slides.mitchelloharawild.com/pruning-graphs/`

2023 **Creating flexible e-learning quizzes with literate programming**: Literate programming Moodle quizzes.
`https://slides.mitchelloharawild.com/moodlequiz/`

2023 **From forecast to fable, design decisions for statistical software**: Design changes from forecast to fable.
`https://slides.mitchelloharawild.com/nyr2023/`

2023 **Reconciliation of structured time series forecasts with graphs**: Coherent forecasting with graphs.
`https://slides.mitchelloharawild.com/reconciling-graphs/`

2023 **The design of statistical software**: Discussion about design principles of user interfaces for statistics.
`https://slides.mitchelloharawild.com/statistical-software-design/`

2020 **Forecasting with multiple seasonality**: Methods and techniques to multiple seasonal forecasting in R
`https://slides.mitchelloharawild.com/nhs2020/`

2020 **Probabilistic cross-temporal hierarchies**: Recent developments of temporal reconciliation in fable.
`https://slides.mitchelloharawild.com/isf2020/`

2019 **Flexible futures for fable functionality**: Introduction to extensible tidy forecasting with fable.
`https://slides.mitchelloharawild.com/fable-tfeam/`

2019 **Flexible futures for fable functionality**: Reconciled forecasting of time series with model combinations.
`https://slides.mitchelloharawild.com/isf2019/`

2018 **Tidy forecasting in R**: Discussion of recent developments to the fable framework.
`https://slides.mitchelloharawild.com/fable-tfeam/`

2018 **Forecasting multiple seasonality with state switching**: The FASSTER model with the fable framework.
`https://slides.mitchelloharawild.com/user2018`

2017 **Models for forecasting multiple seasonality**: An introduction to the FASSTER model.
`https://slides.mitchelloharawild.com/melburn17/`

Publications

1. Athanasopoulos, G., Hyndman, R. J., Kourentzes, N., & O'Hara-Wild, M. (2022). Probabilistic forecasts using expert judgment: The road to recovery from COVID-19. *Journal of Travel Research*. `https://doi.org/10.1177/00472875211059240`

Teaching experience

University teaching

I have 10 years of teaching experience at Monash University, mostly teaching applied forecasting, statistical modelling, and data analysis skills. My teaching has been consistently recognised with positive student evaluations and individual praise from my students, with 13 congratulatory letters for outstanding student evaluations (units in top 5%). I am currently co-lecturing Applied Forecasting and Advanced R Programming.

2024-2025 **ETC4500/ETC5450**: Advanced R Programming (Monash Clayton)
2019-2025 **ETC3550/ETC5550**: Applied Forecasting (Monash Clayton)
2020-2021 **ETC5523**: Communicating with Data (Monash Clayton)
2018-2019 **ETC3580**: Advanced Statistical Modelling (Monash Clayton)
2017-2019 **ETC1010**: Data Modelling and Computing (Monash Clayton)
2016 **ETF2700**: Mathematics for Business (Monash Caulfield)
2016-2018 **ETF3231/ETF5231**: Business Forecasting (Monash Caulfield)

Workshops and trainings

In addition to on-campus teaching, I also teach data analysis workshops to people of all backgrounds around the world.

June 2025 **Exploratory time series analysis**: Instructor (ISF, Online)
Half-day workshop on visually exploring tidy time series data.

May 2025 **Tidy time series analysis and forecasting**: Instructor (SSA, Online)
Half-day workshop on time series forecasting with the fable package.

Oct. 2024 **Advanced R Tips and Tricks**: Instructor (WOMBAT, Australia)
Half-day workshop on advanced R programming.

Oct. 2024 **Interactive web applications with Shiny for R**: Instructor (WOMBAT, Australia)
Half-day workshop on creating interactive web applications in R using Shiny.

Oct. 2023 **Tidy Time Series and Forecasting in R**: Instructor (Dept. Edu, Australia)
Two day workshop on forecasting using tidy forecasting tools in R.

Sept. 2024 **AFRICAST: Tidy forecasting for social good**: Instructor (F4SG, Online)
Five day workshop on forecasting for 181 students from 32 Sub-Saharan African countries.

July 2024 **Tidy time series analysis and forecasting**: Instructor (useR!, Austria)
Quick-start workshop for tidy forecasting in R.

June 2024 **Tidy time series & forecasting in R**: Instructor (UM6P, Morocco)
Two day workshop on forecasting using tidy forecasting tools in R.

Jun. 2024 **Tidy time series analysis and forecasting**: Instructor (R/Medicine, Online)
Quick-start workshop for tidy forecasting in R.

Dec. 2023 **Interactive web applications with Shiny for R**: Instructor (SSA, Online)
3-hour online quick-start workshop for the basics of Shiny.

Oct. 2023 **AFRICAST: Tidy forecasting for social good**: Instructor (F4SG, Online)
Five day workshop on forecasting for 62 students from 9 Sub-Saharan African countries.

July 2023 **Tidy Time Series and Forecasting in R**: Instructor (NYR, USA)
Two day workshop on forecasting using tidy forecasting tools in R.

Dec. 2022 **Interactive web applications with Shiny**: Instructor (WOMBAT, Australia)
Half-day workshop on the basics of shiny applications.

Jan. 2020 **Tidy Time Series and Forecasting in R**: Teaching assistant (rstudio::conf, USA)
Two day workshop with Rob Hyndman on forecasting using tidyverse workflows.

Nov. 2019 **Interactive documents with Shiny**: Instructor (CSIRO, Australia)
Two day intermediate workshop on developing of shiny applications.

Sep. 2019 **Data Wrangling**: Instructor (Monash University, Australia)
A short workshop in the 'R Workshops for Beginners' series on using tidyr and dplyr to wrangle data.

July 2019 **Tidyverse developer day**: Helper (RStudio, France)
A one day developer day where I helped R users resolve issues on tidyverse packages.

Aug. 2019 **High-dimensional time series analysis**: Teaching assistant (ISI WSC, Malaysia)
One day short course with Rob Hyndman for analysing and forecasting large collections of time series.

Mar. 2019 **The grammar of animation**: Instructor (Monash University, Australia)
A short workshop for NUMBAT group on creating animated graphics in R with the gganimate package.

Dec. 2016 **Master R Developer Workshop**: Teaching assistant (RStudio, Australia)
A two day workshop with Hadley Wicham on programming with R and developing R packages.

Supervision and mentoring

2025	**Monash Business Analytics**: Co-supervisor Supervised a research project on data cleaning unvalidated values with LLMs.
2022-2023	**Google Summer of Code**: Co-supervisor Supervised the development of the 'texor' R package for converting LaTeX articles to R Markdown.
2022	**Monash Business Analytics**: Co-supervisor Supervised research into hexagonal binning of spatial geometries (common in election visualisation).
2015-2016	**Access Monash**: Mentor Mentored underprivileged high-school students to support their studies and pursuit of career goals.

Awards & achievements

Awards

2023	Student Presentation Award	*ISF*
2021	Dean's Citation for Outstanding Contribution to Student Learning as a Teaching Associate	*Monash*
2017	Commerce Dean's Honour	*Monash*
2016	Commerce Dean's Commendation	*Monash*
2014-2016	Science Dean's List	*Monash*
2014	International Institute of Forecasters Award	*IIF*
2013	Rotary Youth Leadership Award	*Rotary*

Scholarships

2023-2026	Monash Graduate Excellence Scholarship	*Monash*
2023-2026	Monash Faculty Scholarship	*Monash*
2023-2026	Monash Departmental Scholarship	*Monash*
2017	Econometrics Honours Memorial Scholarship	*Monash*
2015 & 2016	Monash Community Leaders Scholarship	*Monash*
2011 & 2012	Mitcham Rotary Scholarship	*Rotary*

Personal projects

I am passionate about permaculture and sustainability, a proud keeper of bees and chickens, and grower of fruits, vegetables and mushrooms. I blend my expertise in data analysis and engineering with my love of nature to monitor and maintain the ecosystems I create. My latest project is designing networked sensors for wide-area sampling of the local environment, primarily using ecoacoustics to analyse sounds from birds, trees, bees, and soil.