Statistics Research in Aotearoa New Zealand: An International Comparison

---

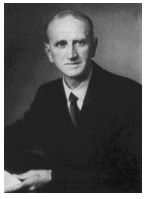
Martin Hazelton <martin.hazelton@otago.ac.nz>
Department of Mathematics & Statistics, University of Otago

NZSA Conference, Dunedin, 8–10 December 2025

New Zealand's Statistical Research Heritage

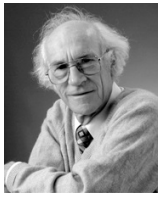
New Zealand has a proud history of world-class statistical research.

Alexander Aitken



Peter Whittle

David Vere-Jones



Shayle Searle

R

From Hero to Zero?

- In the past, NZ has arguably punched well above its weight in statistical research.
- This talk looks at trends over the period 2000–2024.
- Use bibliometrics to compare NZ statistical research against international competitors:
  - Australia
  - Canada
  - Great Britain
  - United States

Journals

- Choice of journals for comparisons based on JUFO Publication Forum (Julkaisufoorumi)
  - Created by Federation of Finnish Learned Societies.
  - JUFO portal: `https://jfp.csc.fi/jufoportal`
- Journals receive rating on 0–3 scale:

3	Top tier – high impact and quality
2	Leading journals
1	Basic level for most peer-reviewed journals
0	Fails to meet criteria for basic level

- Retrieve all rank 2 and 3 journals from Mathematics and Statistics Panel.
  - Maths journals provide comparison.
- Identify by ISSN for future analysis.

JUFO Top Tier (3) Statistics Journals

Annals of Applied Probability
Annals of Applied Statistics
Annals of Probability
Annals of Statistics
Biometrika
Journal of Computational and Graphical Statistics
Journal of the American Statistical Association
Journal of the Royal Statistical Society Series B

JUFO Leading (2) Statistics Journals

Advances in Applied Probability
Bayesian Analysis
Bernoulli
Biometrics
Biostatistics
Electronic Journal of Probability
Electronic Journal of Statistics
Journal of Multivariate Analysis
Journal of the Royal Statistical Society Series A: Statistics in Society
Journal of the Royal Statistical Society Series C: Applied Statistics
Journal of Time Series Analysis
Scandinavian Journal of Statistics
Statistica Sinica
Statistical Science
Statistics and Computing
Statistics in Medicine
Stochastic Processes and Their Applications
Technometrics

Bibliometric Data Sources

Bibliometric data derived from two sources.

OpenAlex `openalex.org`

- Free bibliographic catalogue of scientific papers.
- Equipped with reasonable API (application programming interface).
- Used here to give author country affiliation and publication year (online) for journal articles.

Crossref `www.crossref.org`

- Provides DOI (digital object identifier) information.
- Used here to give number of articles per journal issue.

Useful Packages in R

`httr`

Tools for interacting with web APIs and performing HTTP requests.

`jsonlite`

Tool for working with JSON (JavaScript Object Notation) data.

`rcrossref`

Interface to various Crossref APIs.

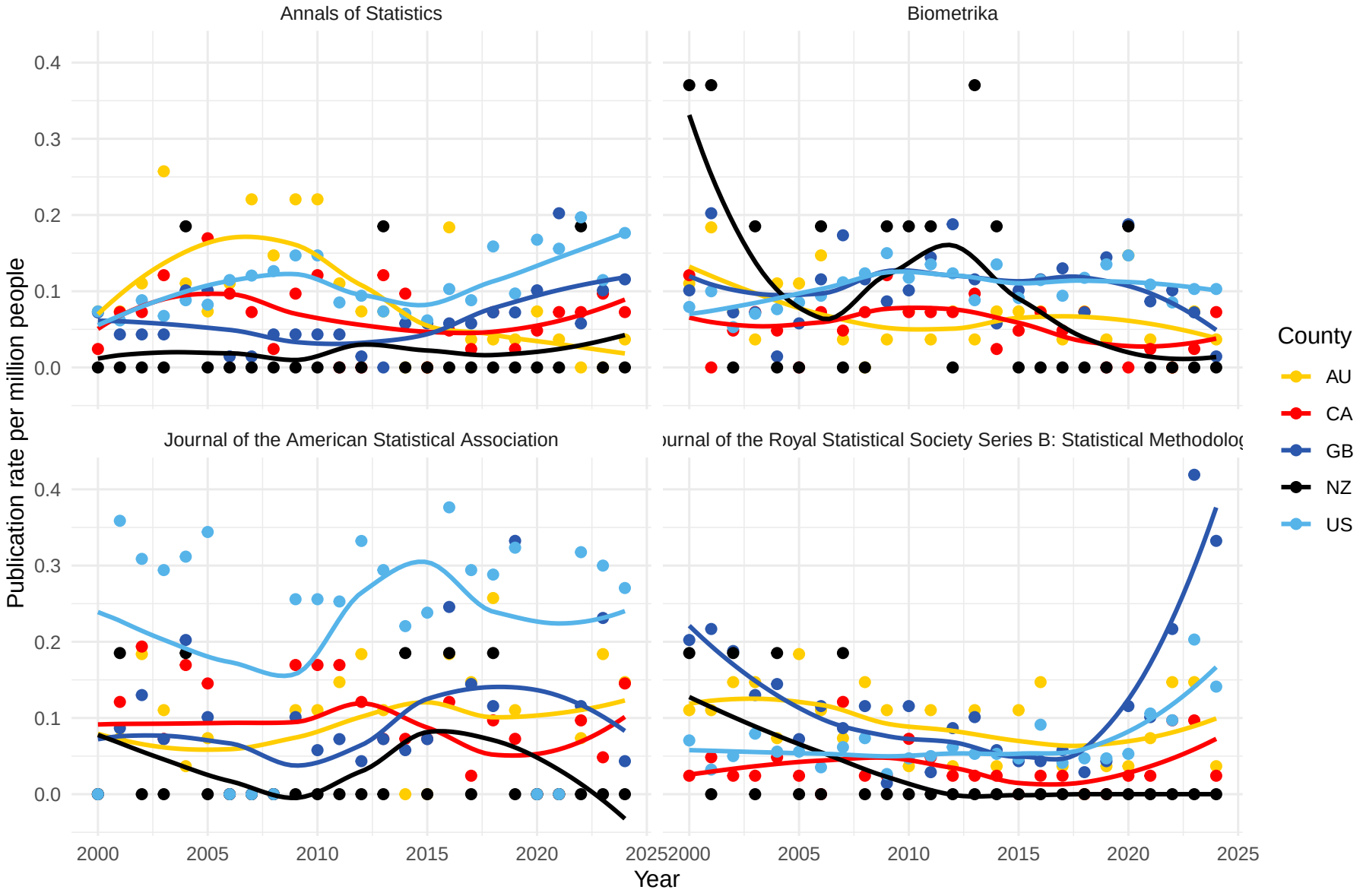
`tidyverse`

Not always a fan of `tidyverse` packages, but convenient here.

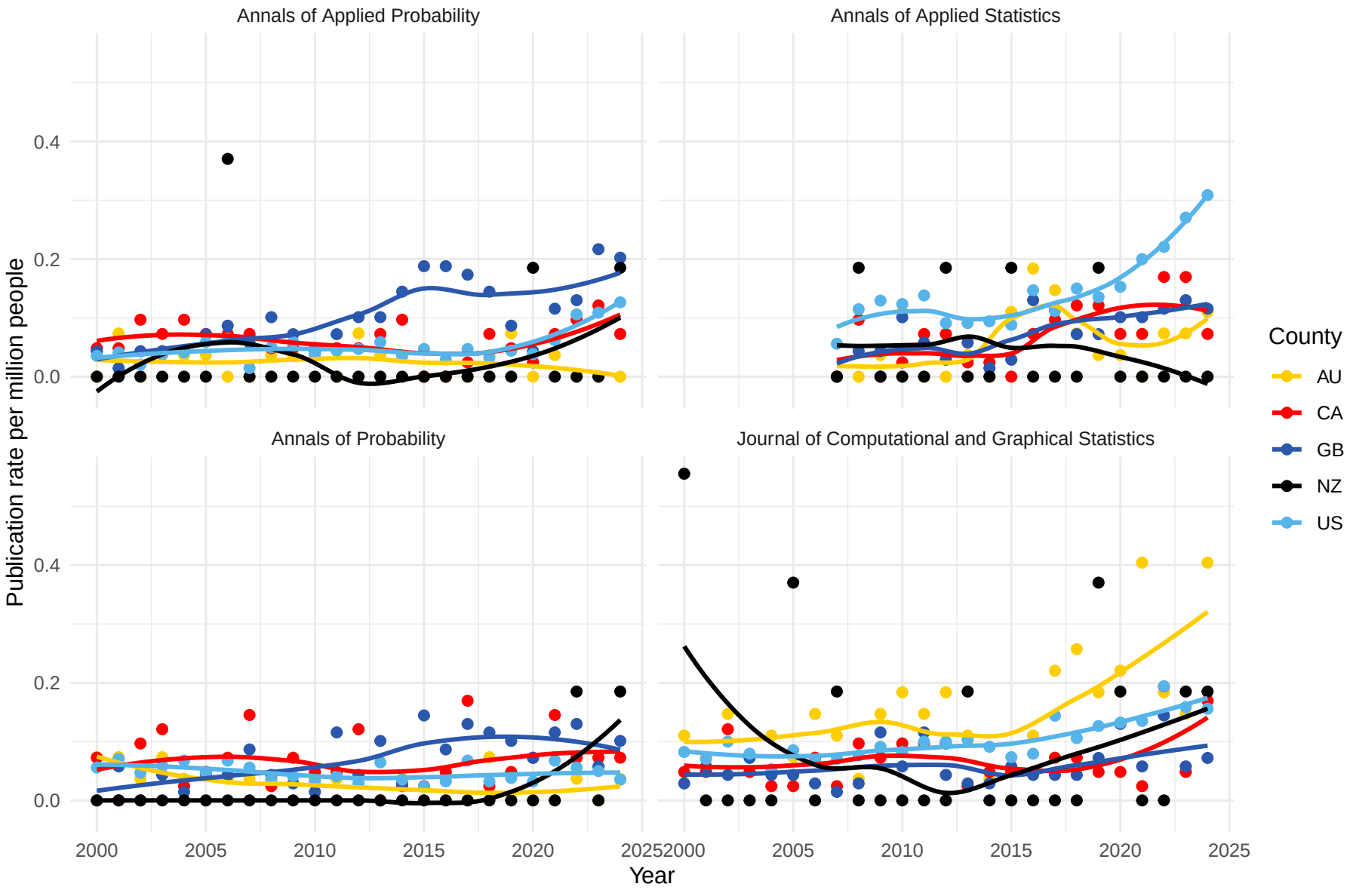
R Data Frame

```
> str(pubs)
'data.frame':   19750 obs. of  11 variables:
 $ ISSN              : chr  "0065-1036" "0065-1036" "0065-1036" "0065-103
 $ year              : num  2000 2000 2000 2000 2000 ...
 $ journal           : chr  "Acta Arithmetica" "Acta Arithmetica" "Acta A
 $ country           : chr  "AU" "CA" "GB" "NZ" ...
 $ JUFO              : num  2 2 2 2 2 2 2 2 2 2 ...
 $ count.all.authors : num  2 4 5 0 13 0 4 4 0 35 ...
 $ count.first.author: num  1 2 3 0 11 0 3 4 0 21 ...
 $ subject           : Factor w/ 2 levels "Mathematics",..: 1 1 1 1 1 1 1
 $ unis              : int  28 20 50 7 126 28 20 50 7 126 ...
 $ population        : num  27.2 41.3 69.2 5.4 340.1 ...
 $ articles          : num  101 101 101 101 101 120 120 120 120 120 ...
```

First Author Publication Rate per Million People



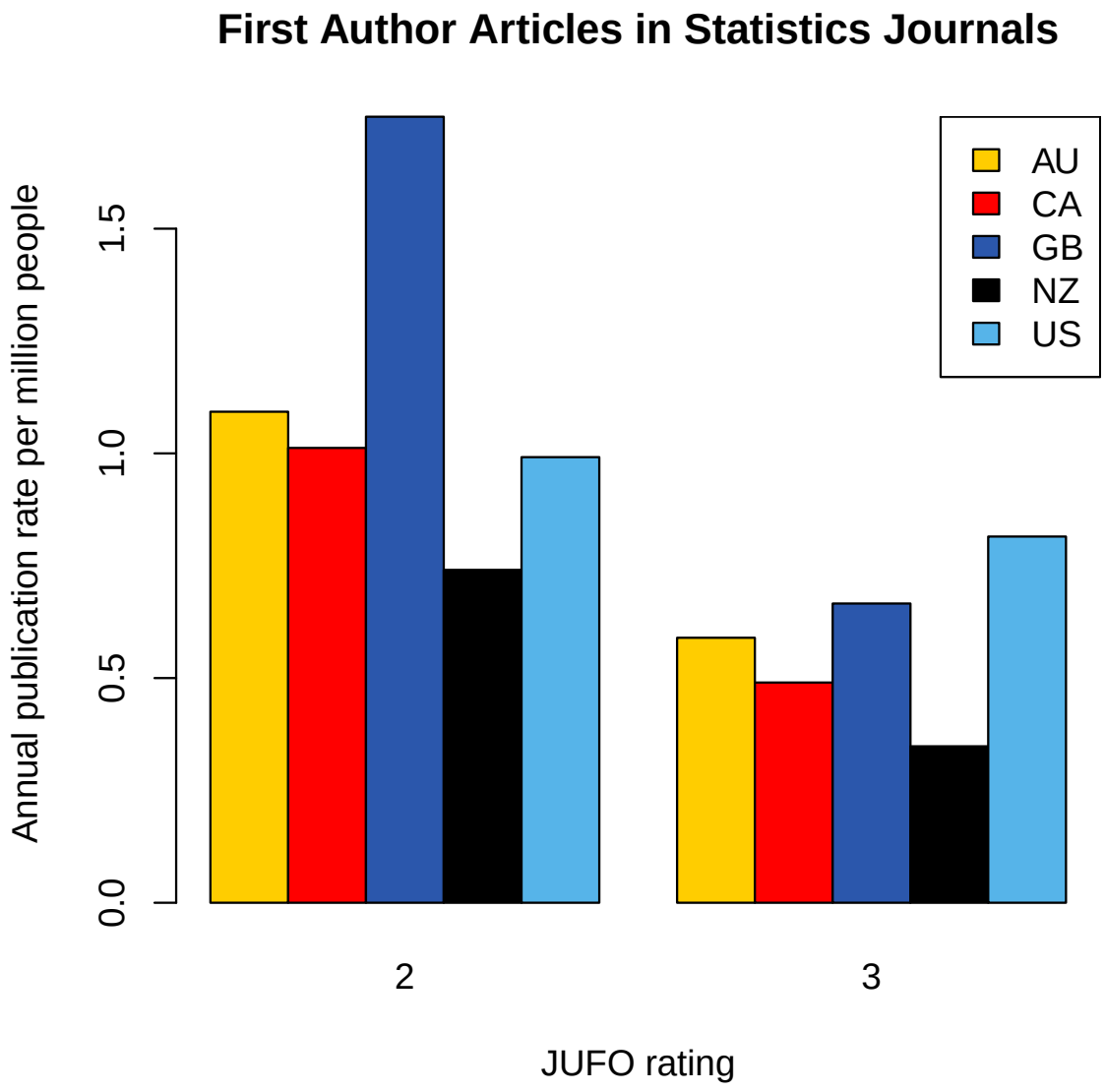
First Author Publication Rate per Million People



Aggregated Data Frames

```
> stats.pubs.agg
   country JUFO count.first.author articles population
1       AU    2                743    45790       27.2
2       CA    2               1045    45790       41.3
3       GB    2               3026    45790       69.2
4       NZ    2                100    45790        5.4
5       US    2               8431    45790      340.1
6       AU    3                401    19162       27.2
7       CA    3                506    19162       41.3
8       GB    3               1152    19162       69.2
9       NZ    3                 47    19162        5.4
10      US    3               6929    19162      340.1
```

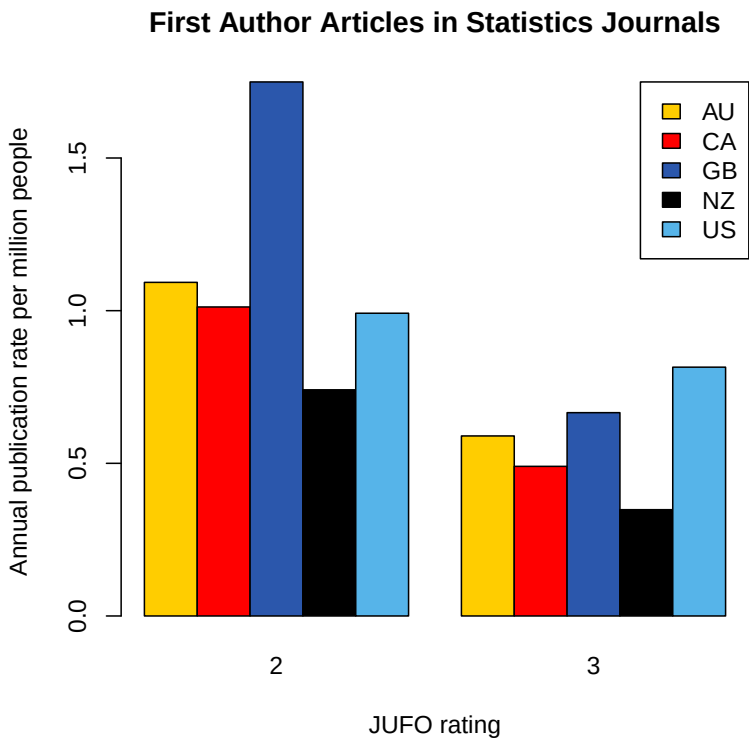
Mean Publication Rates per Million People

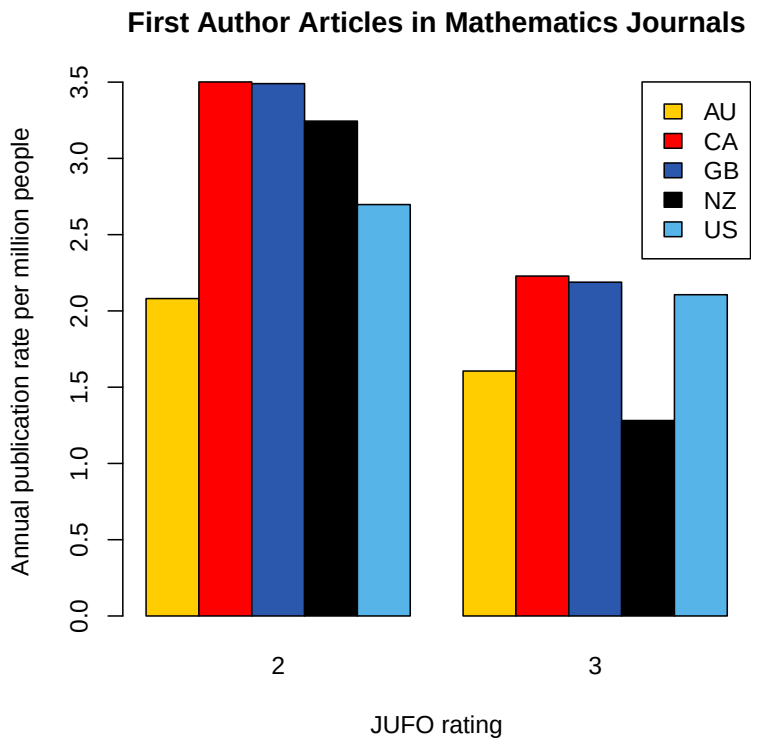


A Simple Model

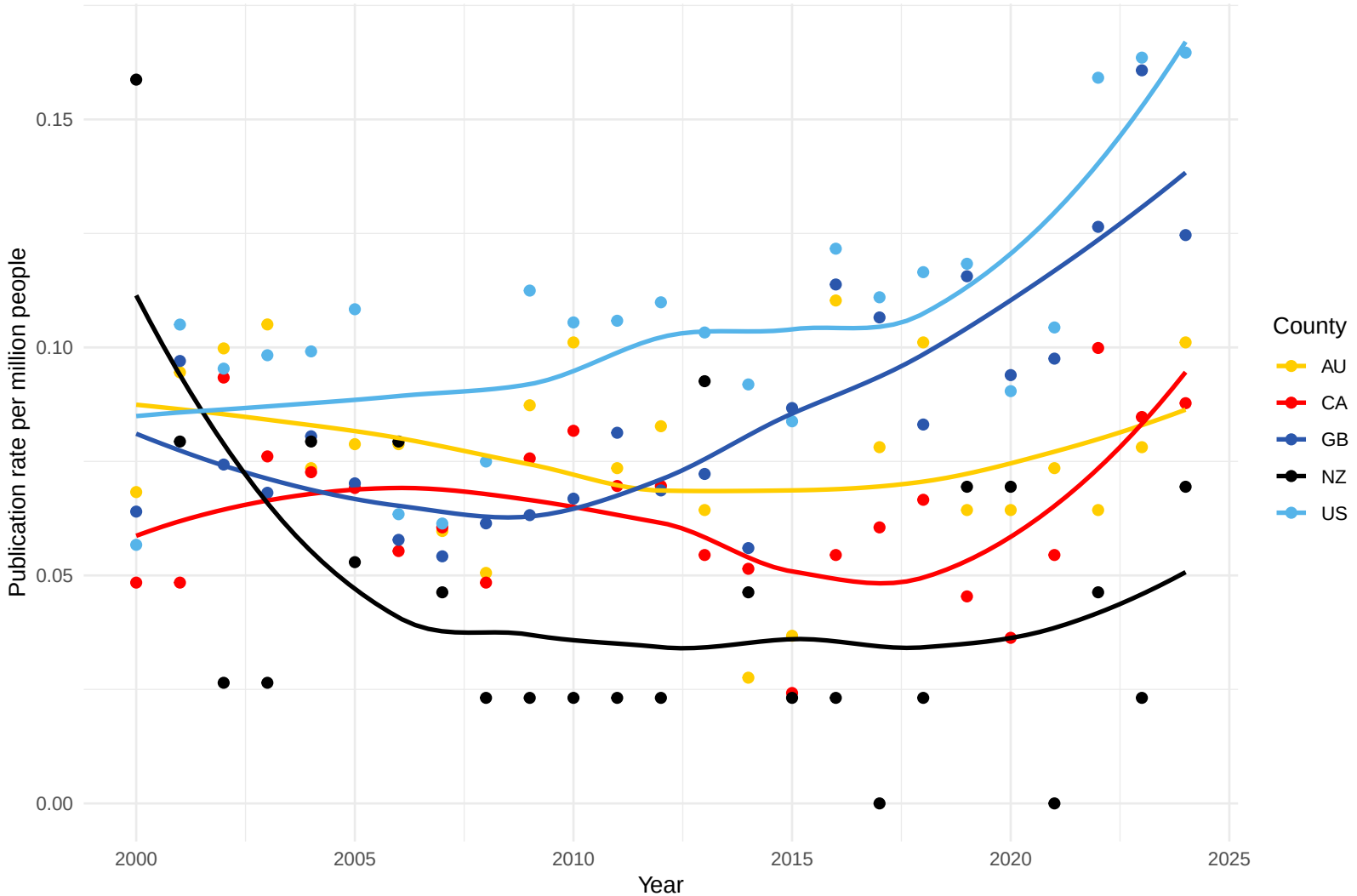
```
> stat.glm <- glm(count.first.author ~ as.factor(country)/as.factor(JUFO),
+         offset=log(population),"poisson",data=stats.pubs.agg)
> summary(stat.glm)

Call:
glm(formula = count.first.author ~ as.factor(country)/as.factor(JUFO),
    family = "poisson", data = stats.pubs.agg, offset = log(population))

Coefficients:
                                      Estimate Std. Error z value Pr(>|z|)
(Intercept)                            3.30748    0.03669  90.155  < 2e-16 ***
as.factor(country)CA                  -0.07657    0.04799  -1.596 0.110578
as.factor(country)GB                   0.47052    0.04094  11.492  < 2e-16 ***
as.factor(country)NZ                  -0.38871    0.10652  -3.649 0.000263 ***
as.factor(country)US                  -0.09705    0.03827  -2.536 0.011214 *
as.factor(country)AU:as.factor(JUFO)3 -0.61673    0.06196  -9.953  < 2e-16 ***
as.factor(country)CA:as.factor(JUFO)3 -0.72524    0.05416 -13.391  < 2e-16 ***
as.factor(country)GB:as.factor(JUFO)3 -0.96574    0.03462 -27.896  < 2e-16 ***
as.factor(country)NZ:as.factor(JUFO)3 -0.75502    0.17685  -4.269 1.96e-05 ***
as.factor(country)US:as.factor(JUFO)3 -0.19620    0.01622 -12.100  < 2e-16 ***
```

Publication Rates: Comparison with Mathematics

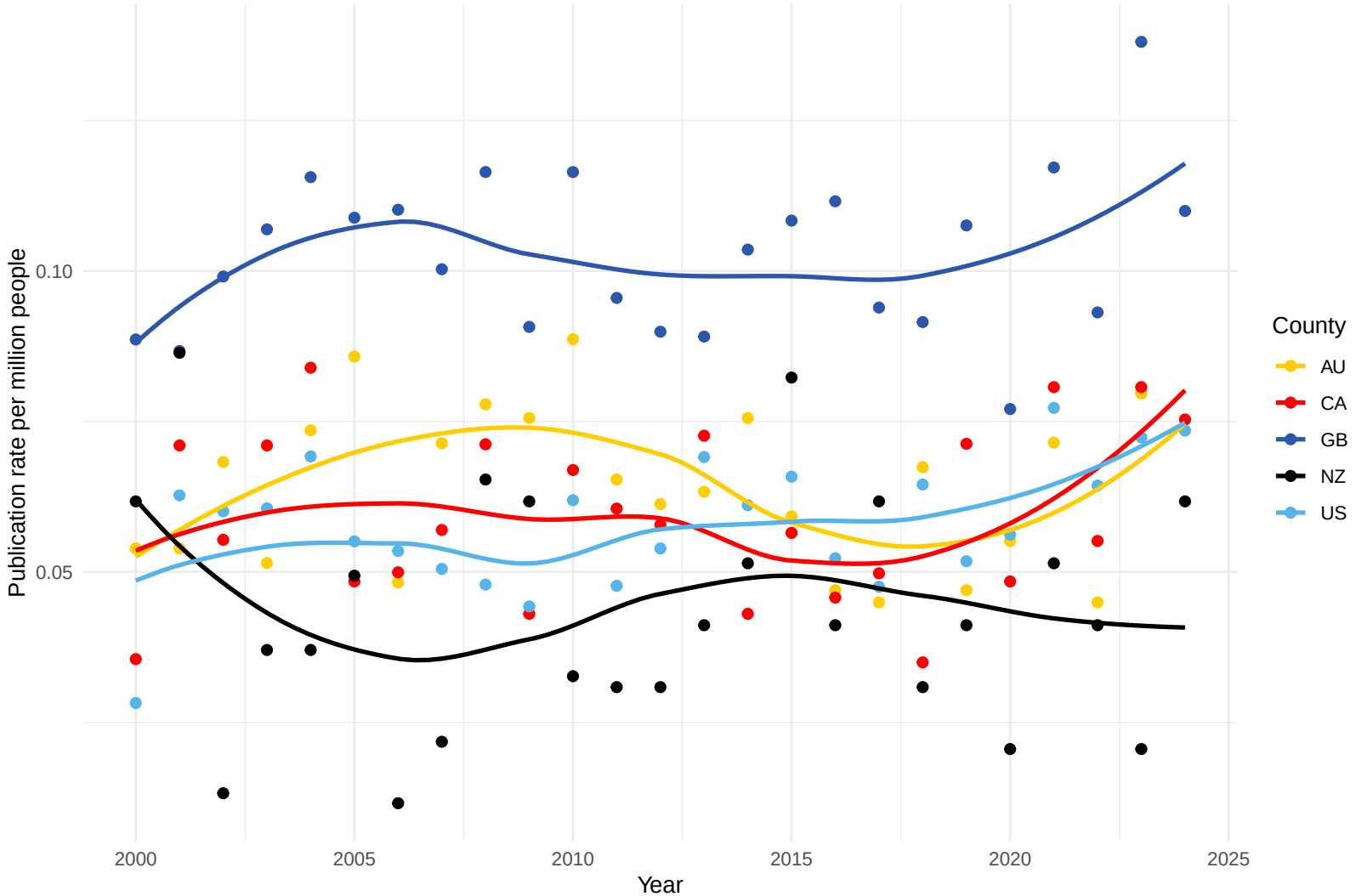




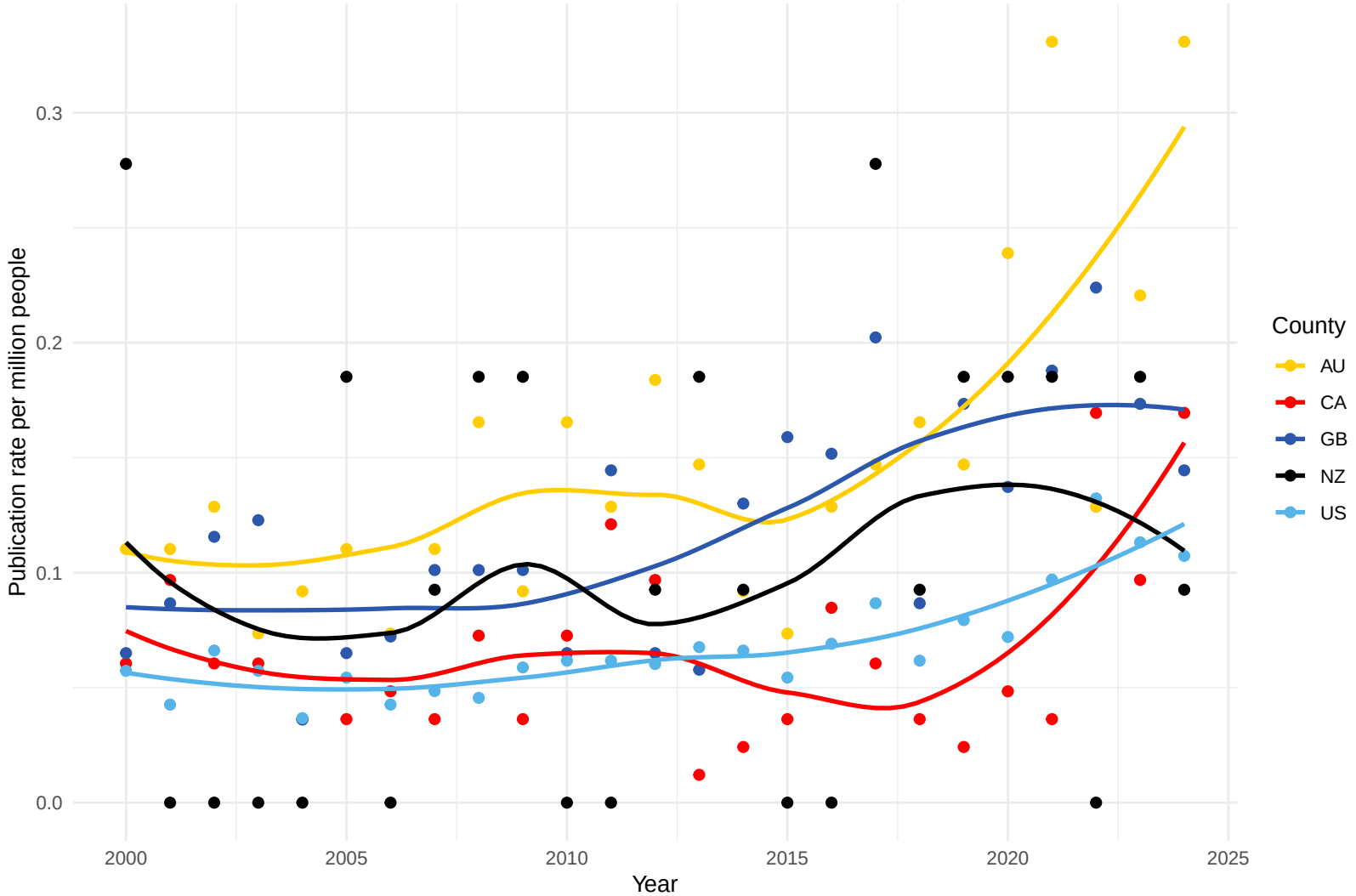
Aggregate Temporal Trends for JUFO 3 Statistics

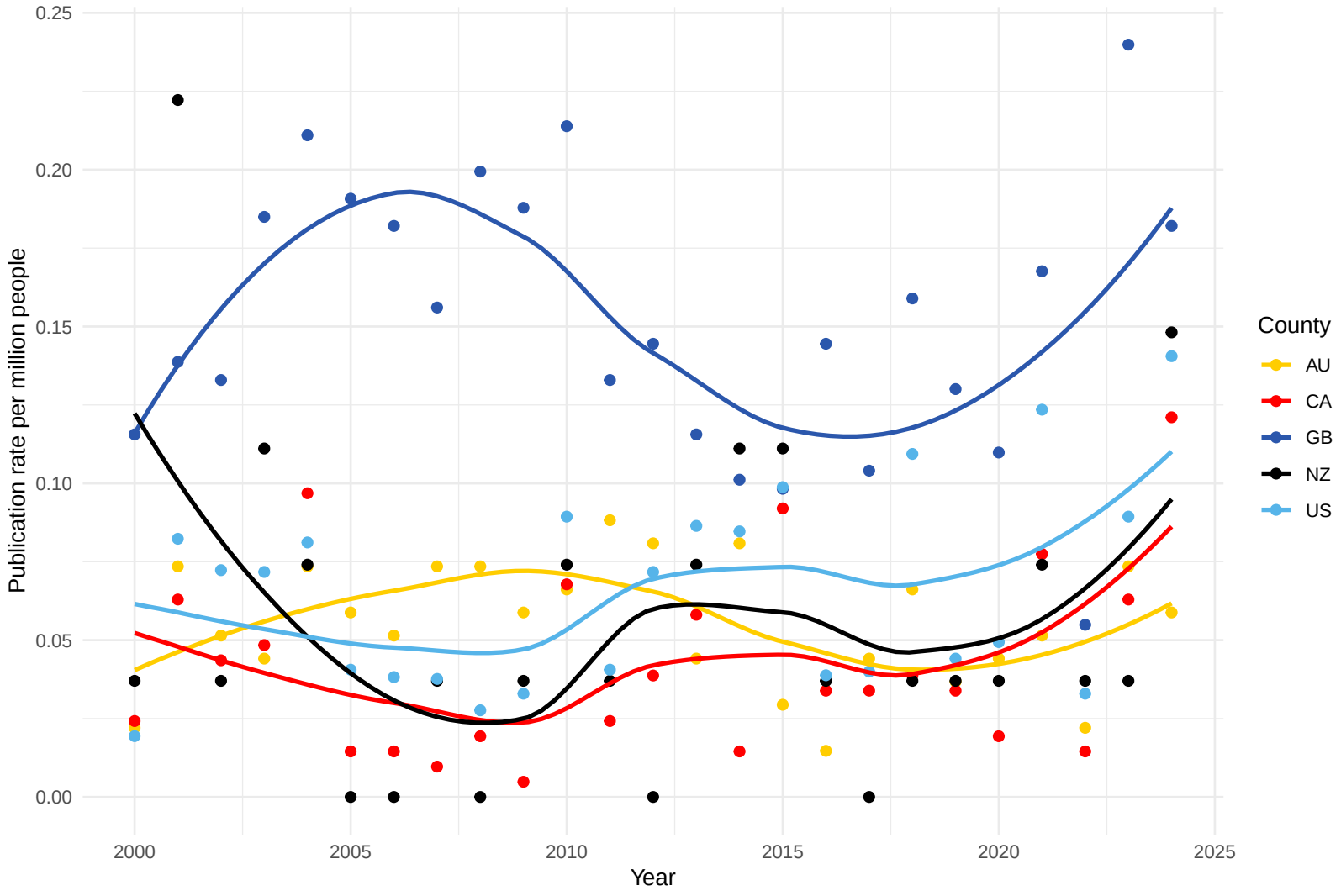


Aggregate Temporal Trends for JUFO 2 Statistics

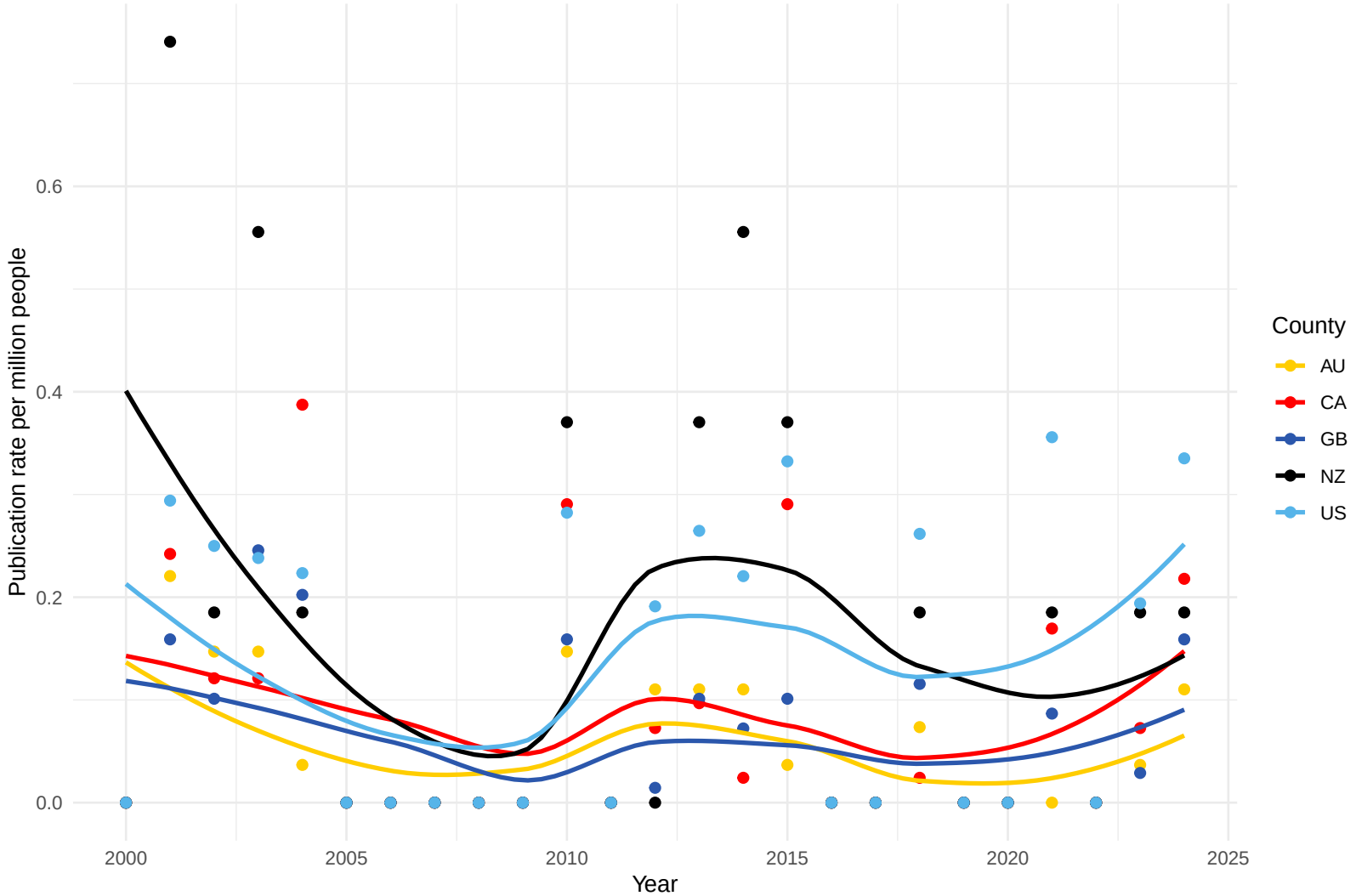


Statistical Computing




0.25
0.20
0.15
0.10
0.05
0.00
Publication rate per million people
2000
2005
2010
2015
2020
2025
Year
County
AU
CA
GB
NZ
US

Biometrics


Publication rate per million people
0.0
0.2
0.4
0.6
2000
2005
2010
2015
2020
2025
Year
County
AU
CA
GB
NZ
US

Summary

- Over the past 25 years, NZ has underperformed on publications in statistical theory and methods in top journals.
- If anything, the temporal trend is downward.
- Remaining steady and competitive in statistical computing, applied statistics, and biostatistcs.
- End on a high note: NZ is comparatively strong in statistical ecology.