Pacing Profiles in World Championship 2000m Rowing:
Explored through k-Shape Clustering

Dani Chu, Ryan Sheehan, Dr. Jack Davis, and Dr. Ming Chang-Tsai

Outline

▸ World Championship 2000m Rowing Data

▸ k-Shape Clustering

▸ Factors associated with Pacing Profiles

▸ Limitations and Future Work

Data

www.worldrowing.com:

- ▸ Olympics
- ▸ Paralympics
- ▸ **World Championships**
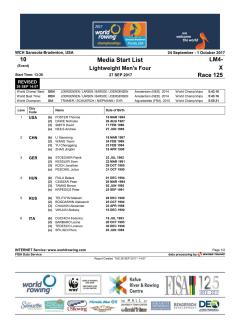- ▸ World Cups

- ▸ Continental
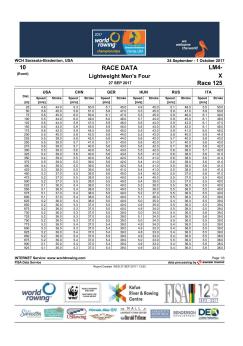- ▸ Under 23
- ▸ Junior
- ▸ Continental

Data: www.worldrowing.com

Media Start List

Race Data

Results

Data: www.worldrowing.com

1. Scrape PDF files from World Championships
2. For each race, extract data from the 3 PDFs
3. Join the race data from the 3 PDFs
4. Combine all races into one file
5. Make code and data available on github!

   github.com/danichusfu/rowing_pacing_profiles

Goals!

▸ Identify the pacing profiles being used by each boat in World Championship 2000m Rowing

▸ Identify which race factors are associated with exhibiting a pacing profile
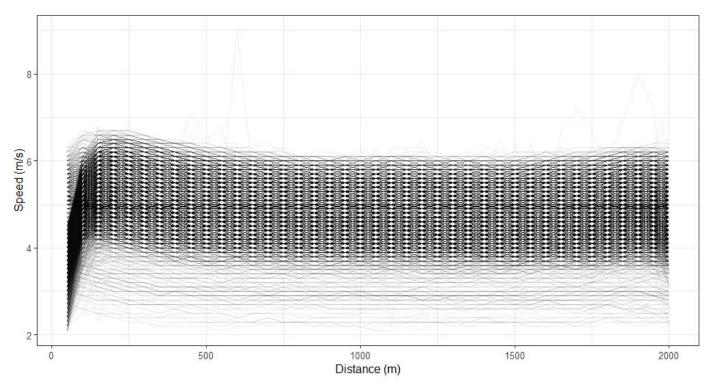
Identification of Pacing Profiles

- Cluster boats based on their average speed at each 50m split
- Problems:
  - Longitudinal data
  - Magnitudes of average speed depend on factors such as boat size, weight class, age group and gender
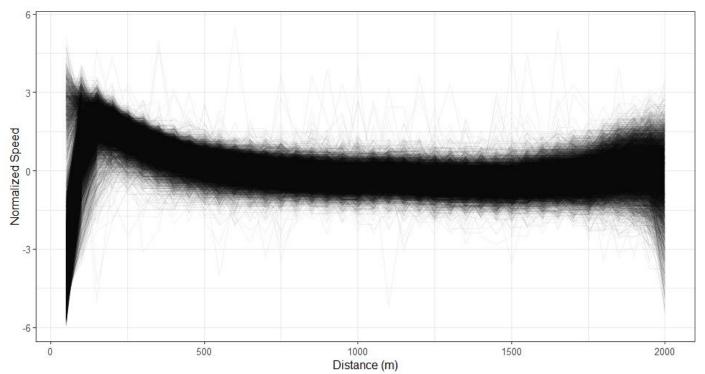
Raw Speed Curves

Normalized Speed Curves
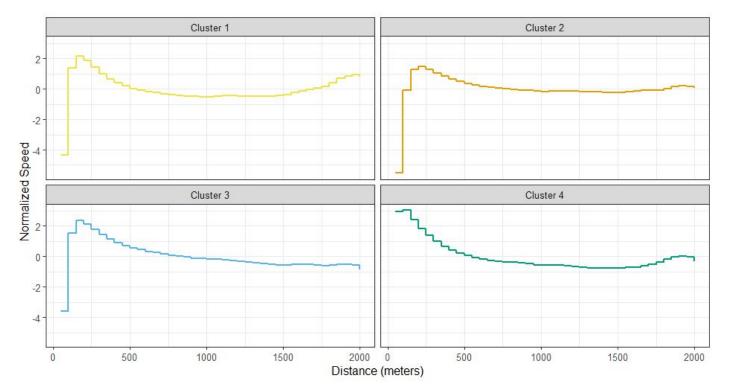
k-Shape Clustering

▸ Uses Shape-based distance (SBD) as an alternative to Dynamic Time Warping (DTW)

▸ SBD is computationally more efficient than DTW

- $O(m \log(m))$ to $O(m^2)$

▸ Small sacrifice in accuracy in experimental settings
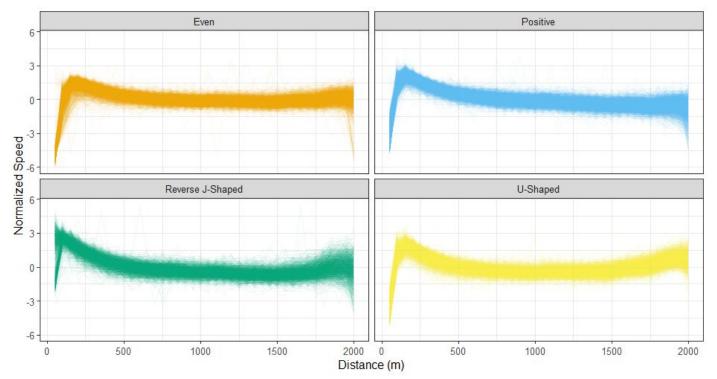
Cluster Centroids

Pacing Profiles

Boat Sizes
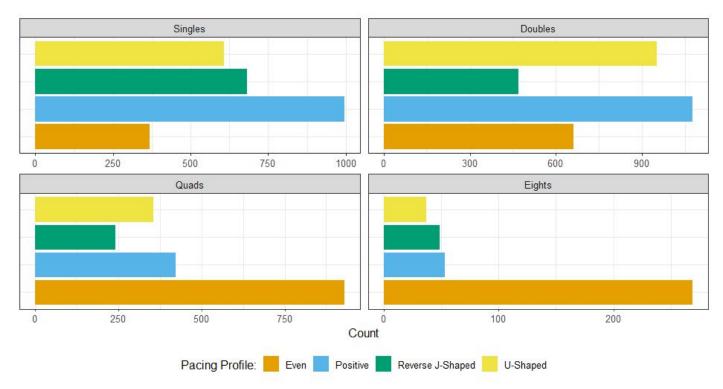

Pacing Profile: Even, Positive, Reverse J-Shaped, U-Shaped

Modelling Pacing Profiles

▸ Response variable is the identified pacing profile

▸ Modelled as a function of race factors

▸ Using multinomial logistic regression

  ● Even profile is the baseline

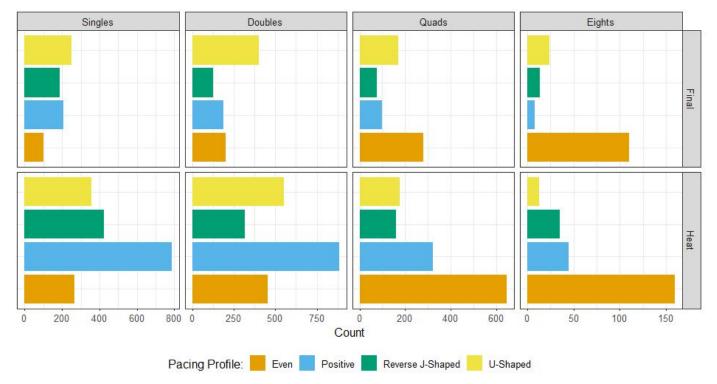 ● Report the relative risk ratio for a one-unit increase in the variable

Model Results

	Positive	Reverse J-Shaped	U-Shaped
Intercept	0.90	0.77	1.41
Size: Doubles	**0.48**	**0.38**	**0.68**
Size: Quads	**0.13**	**0.14**	**0.16**
Size: Eights	**0.04**	**0.08**	**0.04**
Heat or Final: Heat	**1.81**	1.04	**0.54**
Race Placement: 2nd Place	0.86	1.02	1.21
Race Placement: 3rd Place	1.08	1.33	**1.50**
Race Placement: 4th Place	1.36	**1.60**	**1.60**
Race Placement: 5th Place	**1.76**	**1.93**	1.24
Race Placement: 6th Place	**3.16**	**3.16**	1.21
Discipline: Sweep	**1.81**	1.20	**1.97**
Gender: Women	**1.88**	**1.68**	**1.66**
Weight Class: Open	**1.43**	**1.52**	**1.28**

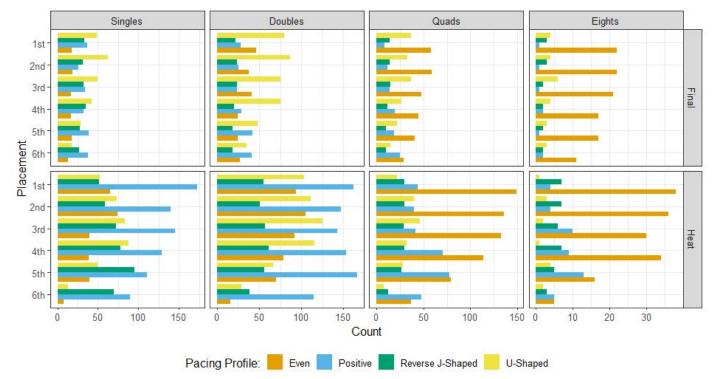# Boat Sizes and Heat/Final

Boat Sizes, Heat/Final and Placement

Conclusions

▸ Can identify pacing profiles with k-Shape Clustering

▸ Interesting preliminary results for which race factors affect pacing profiles

▸ Available data: github.com/danichusfu/rowing_pacing_profiles

▸ I'd love your feedback and thoughts!

Limitations

- ▸ Observational Data

- ▸ Interaction terms are not fit in the model

- ▸ Cannot choose a "optimal" profile to help coaches and athletes

- ▸ Only uses World Championship Races

THANKS!

Any Questions?

Dani Chu

dani_chu@sfu.ca

@chuurveg

danichusfu.github.io

Ryan Sheehan

ryan_sheehan@sfu.ca

Dr. Jack Davis

jack_davis@sfu.ca

@jack_davis_sfu

sfu.ca/~jackd/

Dr. Ming Chang-Tsai

mtsai@csipacific.ca

@ming_chang_tsai