Data-driven method for navigating the Atlantic in a rowing race

Hartstein, Máté
mate.hartstein@arcablanca.com

Čevora, George
george.cevora@arcablanca.com

Arca Blanca Ltd.
44 Maiden Lane, London WC2E 7LN

Abstract

Ocean rowing not only features epic distances, but also requires adaption to harsh weather conditions. This makes it essential that rowing crews racing across the ocean exploit wind to aid their success. However, the favourable winds rarely follow the shortest path between the start and finishing line of a rowing race. This common discrepancy requires boat crews to make a compromise between shortest path to victory and the path that exploits the wind. These decisions are currently based on experience, consensus and gut-feel; methods that rarely select truly optimal solution. In this paper we demonstrate a generic data-driven method for finding such optimal compromise by considering various route choices for the Atlantic Rowing Race, which requires rowers to make a crossing from the Canary Islands to Antigua. We conclude that the most common routes on this crossing are biased too far south towards areas with favourable *trade winds* making the journey unnecessarily long. Finally we provide map of a route for this crossing which is optimal according to 2000-2019 weather data.

The Atlantic Rowing Race [1], also known as Talisker Whisky Atlantic Challenge [6] since 2011, is a trans-Atlantic endurance rowing race from the Canary Islands to Antigua. The 4700 km race is unsupported, meaning that each crew has to carry all supplies in their boat. Boats entering this race have to comply with R45 design [2]. While these boats do not have any features specifically designed to utilise wind, just like any floating object with a surface exposed to winds, they are to some degree propelled by wind. This is perhaps best demonstrated by Jean-Jacques Savin, a man who in 2019 crossed the Atlantic in a barrel with no propulsion whatsoever at speed greatly exceeding the North Atlantic Equatorial Current [3]. Wind therefore seems as the main factor with potential to speed up the rowing boats crossing the Atlantic.

We were approached by one of the competing teams [5] seeking advice on how to best use wind in the 2021 race. Three issues have a crucial impact on any viable solution: 1) rowing boats move relatively slowly ($<$8 km/h); 2) the route is very long ($>$4000 km); and 3) wind forecasts become unreliable beyond 5-7 days into the future. Taken together, these three issues mean that it is not possible to navigate the rowing boat according to weather forecasts as the wind is going to significantly diverge from forecasts by the point the boat reaches the desired position. We have therefore defaulted to climatology to select the best route.

Trade winds [7], are wind patterns exploited for a faster crossing of the Atlantic for centuries. However, these winds do not follow the shortest path between the Canary Islands and Antigua, posing a difficult choice for all seamen: follow the shortest path or the fastest wind? While the advantage to be gained from trade winds has not been to our knowledge properly quantified, we are anecdotally aware that essentially all rowing boats attempting the crossing diverge from the shortest path to south in chase of better winds. This is supported by the 6 crossings covered in our dataset (see Fig. 3a). Interestingly, the crews diverging the least from the shortest path (The Four Oarsmen and Team Antigua) managed to achieve the two fastest crossings in the history of the race. This situation begs the question, how much exactly should the crew of a rowing boat prioritise chasing better winds as opposed to pursuing the shortest path to the destination; a question we've chosen to answer.

Only one published attempt at the route optimisation for a transatlantic race is known to us [8]. We improve on this work by 1) providing full transparency regarding our approach, 2) disclosing accuracy of our model, 3) incorporating modern wind data unavailable to Leyland and Philpott, and 4) demonstrating the impact of year-to-year variation in weather on the optimal routing.

1 Methods

Our aim is to devise a model that relates wind conditions (wind speed, and wind direction relative to boat heading) to the expected rowing speed of the boat (hereinafter referred to as the *boat model*) illustrated in Fig. 1. At the outset it is difficult to theorise what form the relationship would take. Drag force is ordinarily considered to scale quadratically with wind speed, but wave height, which hinders rowing, also scales quadratically. With growing wave height, rowing becomes increasingly technical, especially when the waves arrive perpendicular to the boat. For strong enough winds its positive effects are thwarted by the waves, and can even make it impossible to progress.

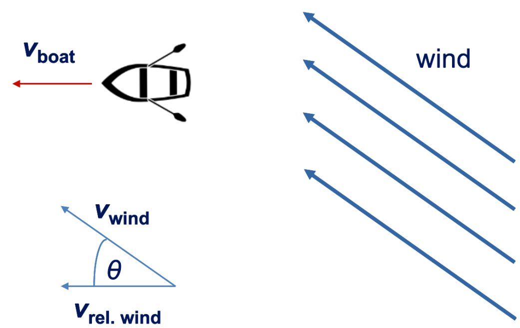

Figure 1: **Schematic diagram showing the rowing boat relative to the wind.** The diagram relates the direction of the wind to the heading of the boat. The relative wind speed ($v_{\mathbf{rel.wind}}$) is defined as the component of the wind speed ($v_{\mathbf{wind}}$) that is parallel with the boat heading.

For wind data in the Atlantic we used the NCEP/NCAR dataset published by NOAA [4], which goes back to 1948. It contains 4-times daily surface wind speeds with a spatial resolution of 250 km. The data are separated into U-wind and V-wind, which are the meteorological wind components referring to the component blowing to the East and the component blowing to the North, respectively, and are given in units of m/s.

We used data from past competitions to obtain actual boat speeds realised in a competition setting. We had access to the paths (GPS coordinates and time) of the fastest two competitors from each of the 2017, 2018 and 2019 Atlantic Rowing Races, but not the synchronous wind speed and boat speed. The locations of each boat were recorder every four hours. All boats were of the R45 boat design [2], meant for a crew of four. We inferred the realised boat speeds (averaged over 4 hours) by considering the distance and elapsed time between two successive points along the paths. The boat heading was defined by the vector formed by successive points along the path. For our analysis we decided to exclude the first segment of the races, where the boats are not settled into a rhythm yet, and the final segment, where the competitors usually go full out, as we found these sections do not provide representative data points. This meant that we only considered the paths that fell between a longitude of 20°W and 60°W.

Finally, we combined the two datasets by building a function that retrieves the wind data given a set of coordinates and a date as defined by the points along the paths of the competitors. We also re-defined the wind angle relative to the heading of the boat (0° for tailwinds and 180° for headwinds).

2 Results

Fig. 2 shows the realised boat speeds (v_{boat}) for six different boats as a function of relative wind speed ($v_{\mathrm{rel.wind}}$). The data for different boats overlap well, demonstrating that the performance of the fastest crews from these three years were comparable. The majority of the data points correspond to positive relative wind speeds, so favourable winds, which is expected for the route. For our boat model we use a simple linear fit (red dashed line) to avoid overfitting. With the noise in the data it was not possible to separate out the opposing effects of the wind drag and wave height, which are both expected to be superlinear. Instead, they appear to combine into a linear relationship for the relative wind speed range under consideration. A decline in boat speed is expected past this range, so above 15 m/s, where wave heights normally exceed 5 m. In the negative relative wind speeds region (headwinds) the two superlinear effects should add up and affect performance negatively, leading to divergence from linearity, but this is not seen conclusively for the limited data we have. We also note the data points for Fortitude IV and Rowed Less Travelled (the two fastest boats from the 2019 race), which show high boat speeds in headwind. These were achieved for a 250 km long section of their paths (longitudes between 29.5°W and 32°W). It is likely the NOAA wind data did not accurately reflect the wind experienced by the boats for this section.

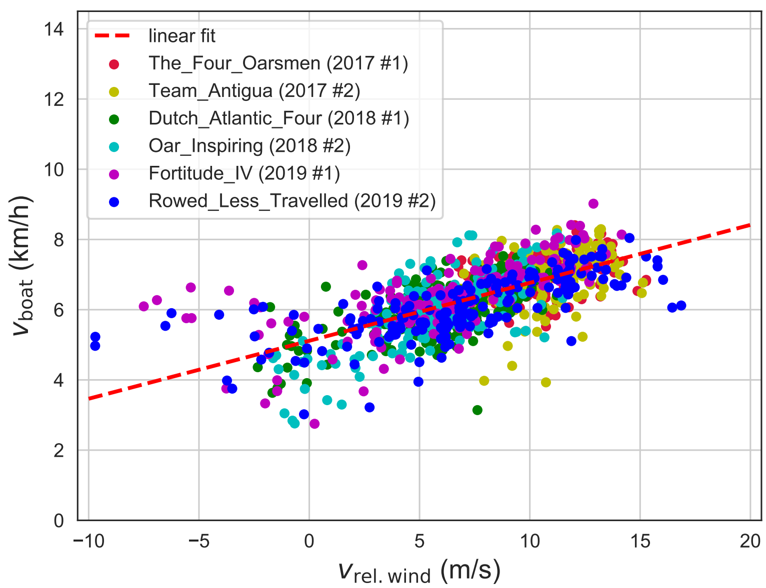

Figure 2: **Boat speed as a function of relative wind speed based on data from past Atlantic crossings.** Boat speed (v_{boat}) values (circles) of the fastest two boats from the Atlantic Rowing Races in 2017, 2018 and 2019 as a function of the relative wind speed ($v_{\text{rel.wind}}$) that corresponds to the same location and time as the boat speed reading (see Methods). A clear correlation is seen regardless of the boat. We perform a linear fit (red dashed line) on the combined dataset to determine the relationship between boat speed and relative wind speed which we use for our simulated Atlantic crossings.

We also endeavoured to determine the polar diagram of the R45 rowing boat. Polar diagrams describe the expected boat speed for a range of wind speeds and relative wind angles, and are commonly used as input for sailing boat route planning. Unfortunately, due to the predominant tailwinds for this route, the data do not cover enough relative wind angles to allow for the construction of the polar diagram for the rowing boat.

To test our boat model, we performed simulations of past crossings (Fig. 3a). For each crossing we considered the initial location and time, and the corresponding wind speed and the wind angle relative to the direction of travel. We input these values into our boat model to calculate the expected boat speed, and using resulting time for the boat to reach the next checkpoint along the path. We simulated the entire path by repeating these steps, and adding up each time segment predicted by the model, until we get to the end of the path. This resulted in the predicted time it would take our simulated boat to travel along the path taken by real boats before.

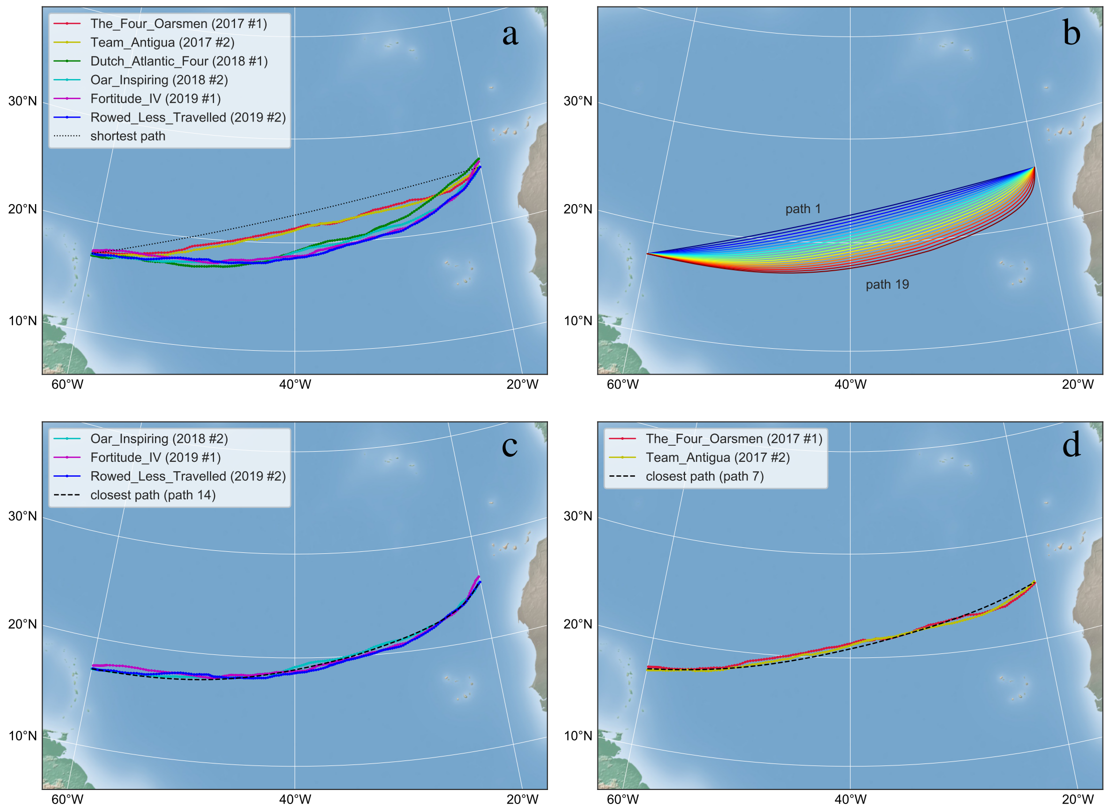

Figure 3: **Comparison of different paths for Atlantic crossings.** a, The paths taken by the fastest two boats from the races in 2017, 2018 and 2019, shown for longitudes between 20°W and 60°W. None of the boats took the shortest, most direct path, but instead went south for more favourable winds. b, We created 19 paths that mimic the paths taken by competitors with the variable being how far south they go. Path 1 (blue line) is the most direct path, and path 19 (brown line) is the most southerly path. c, Comparison of the paths taken by three competitors, that took southerly routes, to path 14, the most similar mimicked path to theirs (black dashed line). d, Comparison of the paths taken by the fastest two competitors, which took more direct paths, to path 7, the most similar mimicked path to theirs (black dashed line).

With this method we simulated for the fastest two crossings from each of the races in 2017, 2018 and 2019, by having our model boat set off on the corresponding starting dates and travel along the same paths as the competitors. The results are shown in Table. 4. We have found that the simulations using our boat model predicted crossing times with a mean error of 0.63 days. The mean error is improved slightly to 0.54 days if we do a quadratic fit to the data shown Fig. 2, but runs the risk of overfitting. Competitors generally finish days apart, so the error of less than a day for the linear boat model is sufficiently accurate, especially if we consider that we only relied on the effect of the wind, and assumed that all boats and crews are the same. Ultimately, by using a limited number of crossings we were able to propose and validate a model, that we can now reliable use to simulate other crossings, crossings with different paths, and different starting dates, including crossings that no crew has actually rowed before.

	Path length (km)	Crossing time (days)	Predicted time by model (days)
The Four Oarsmen (2017 #1)	4276	25.42	26.19
Team Antigua (2017 #2)	4271	25.92	26.16
Dutch Atlantic Four (2018 #1)	4396	30.75	30.32
Oar Inspiring (2018 #2)	4367	30.75	30.18
Fortitude IV (2019 #1)	4424	28.67	30.06
Rowed Less Travelled (2019 #2)	4389	30.17	29.79

Figure 4: **Comparison of actual Atlantic crossing times and predicted crossing times.** We simulated the paths between longitudes 20°W and 60°W taken by the fastest two boats from the races in 2017, 2018 and 2019, but using our model boat, with its boat polars determined in Fig. 2. The crossing times calculated for the simulated Atlantic crossings are found to predict actual crossing times with a mean error of 0.6 days.

To produce actionable insight for 2020 race we have simulated the selected paths for years 2000 to 2019 as shown in Fig. 5. The path number 9 was identified as one with the lowest median crossing. We provide the detailed map of the path in Fig. 6.

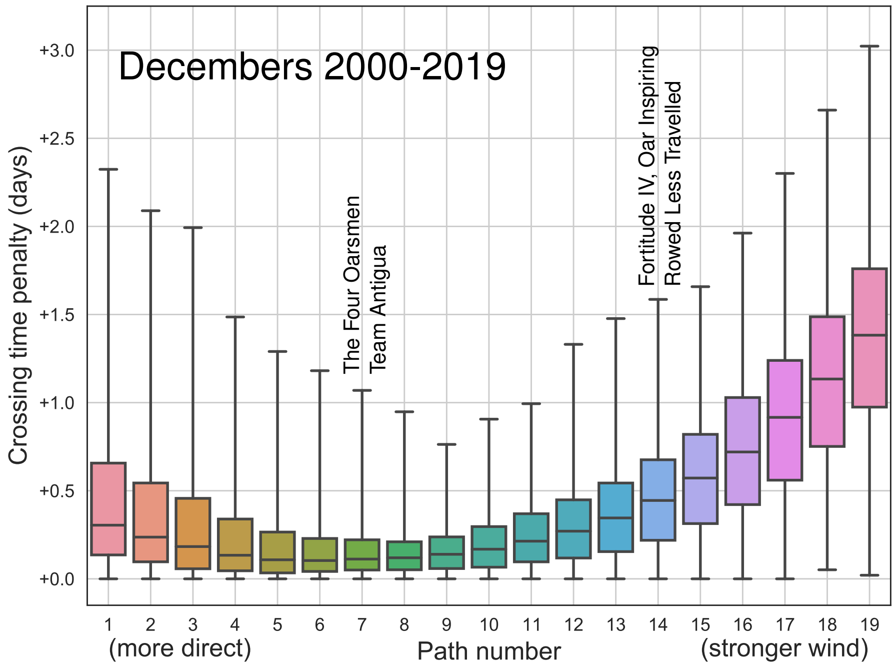

Figure 5: **The aggregated crossing time penalties for different paths.** We simulated Atlantic crossings along the 19 paths shown in Fig. 3b for 600 starting dates covering all Decembers between 2000 and 2019. The figure shows the crossing time penalty for each of the 19 paths. The crossing time penalty is defined as the relative crossing time compared to the fastest path for the same starting date. The boxes represent the interquartile range with the horizontal line corresponding to the median value, and the whiskers mark the entire range of values. We see the lowest median crossing times for paths 5-9, with path 9 being the most consistent. For paths that go further south (more wind) we see an increase in both the median crossing times and the range of crossings time values, indicating that on average they lead to slower, weather-dependent crossings. For paths that are more direct, we only see a small increase in median crossing times, but a similarly large increase in the range of crossing time values, suggesting they are also weather-dependent.

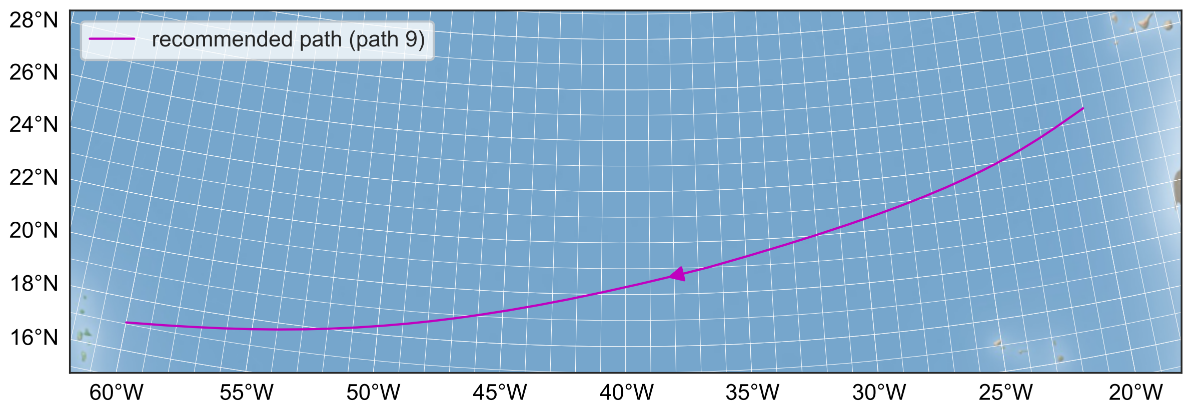

Figure 6: **Our recommended path based on simulations covering 20 years.** Path 9 plotted between longitudes 20°W and 60°W. Based on Fig. 5 path 9 was identified to have one of the lowest median values, and simulations along this path were guaranteed to finish within 0.8 days of the fastest path, making it the most consistent choice.

On Shoulders of Giants, the crew we have supported through the 2020 Talisker Challenge and the winner of the fours category that year, took a route with a total distance of 4385 km (most similar to path 14 of the mimicked paths). Model predicted time based on their actual route which has somewhat diverged from our recommendation was 29.93 days, whereas actual time was 31.14 days. Most of the error (0.67 days) was picked up between 40.3°W and 41°W (100 km path), which is where a storm swept the boat. This truly experimental model validation (if only a singular data point) is well in-line with the error reported in above.

Post-hoc analysis shown in Fig. 7 of weather data from December 2020 reveals that OSOG did not take the optimal path[1], with the optimal route for 2020 being broadly consistent with our general recommendations based on weather between years 2000 and 2019.

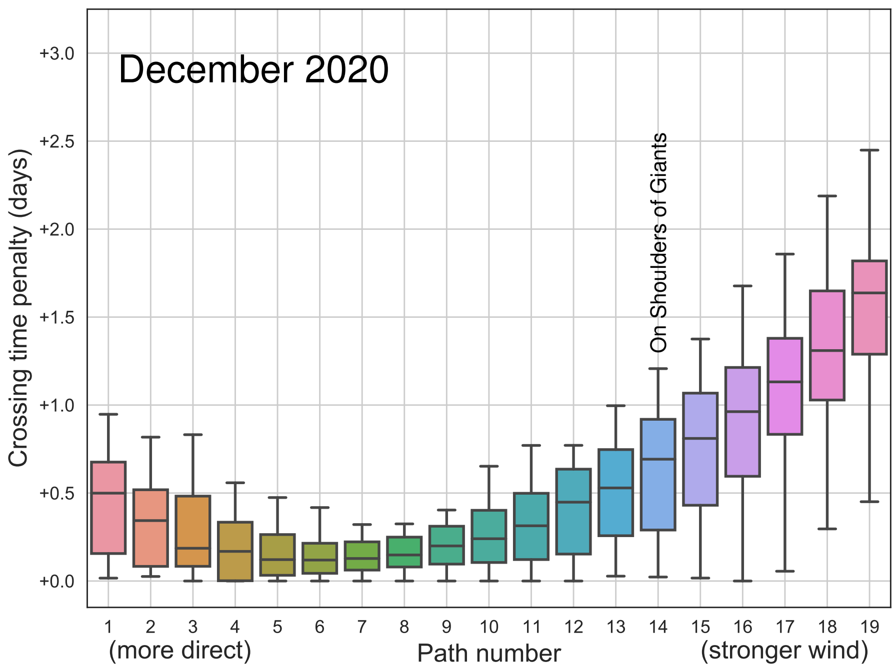

Figure 7: **The aggregated crossing time penalties for December 2020.** We simulated Atlantic crossings along the 19 paths shown in Fig. 3b for starting dates between 1 and 30 December 2020. The figure shows the crossing time penalty for each of the 19 paths. We see the lowest median crossing times for paths 4-8, with paths 7 and 8 being the most consistent. More direct paths were more favourable for this year. On Shoulders of Giants took a path closest to path 14.

[1]This was mostly caused due to unusual danger from storms along the optimal route which made the team divert from recommended route. Storms reaching as far south as during December 2020 have not been recorded during previous 20 years.

3 Discussion

Our paper describes development of a method suitable for selecting the optimal path for a rowing boat across across the Atlantic starting in the Canary Islands and finishing in Antigua. The results we present here focus solely on crossing the Atlantic in a R45 boat design due to the purpose of the project being to support one of the teams competing in the Talisker Whisky Atlantic Rowing Challenge. However, the method we propose is easy to generalize to different crossings of large bodies of water with significant wind and different boat designs depending on the availability of data.

The difficulty in selecting the optimal path lies in the fact that there is a tension between selecting the shortest path and path that utilizes the Trade Winds in an optimal fashion. An appropriate compromise between the two paths should be considered by any crew attempting the crossing, yet no structured method for selecting the optimal path has been published to date. Philpott and Leyland (2005) have devised such a method; however, their publication lacks detail necessary for replication, validation or even application of the method they proposed.

Due to the order of magnitude difference between the daily range of the boat and the coarseness and limited forecast ability of wind data we detailed above, it is impractical to conduct live-routing of any crossing attempts. Instead, we have defaulted to climatology as a source of our information about the wind patterns, which qualifies the recommendations presented here to be applicable to future crossings as long as they take place in December - January: the most favourable season for attempting this feat. The year-to-year variability of optimal path is also evaluated to the conclusion that it makes very small difference.

Four out of six historic crossings we obtained data for have taken crossing routes far southern to the identified optimal paths in chase of better wind, penalising themselves by rowing disproportionately longer distance.

After the conclusion of the 2020 Talisker Challenge we have used the GPS data from the team we have supported, simulated the actual crossing with our model and found that the crossing-time prediction error is well within the range we have anticipated before the crossing took place. While this provides some validation of our model it is a single data point. This situation is emblematic of the main problem in data-driven optimisation for ocean rowing: the scarcity of data.

Improving significantly on the previous attempt at routing a rowing boat across the Atlantic by Philpott and Leyland (2005) we hope this report will provide a springboard for anyone who wishes to optimise the routing even further and also provide a concrete and applicable advice for any crew looking to undertake a row from the Canary Islands to Antigua during the months of December and January.

References

[1] Atlantic Rowing Race. https://en.wikipedia.org/wiki/Atlantic_Rowing_Race. Accessed: 6. 5. 2021.

[2] Boat Details | Talisker Whisky Atlantic Challenge. https://www.taliskerwhiskyatlanticchallenge.com/boat-details/.

[3] Jean-Jacques Savin: Frenchman completes Atlantic crossing in barrel. https://www.bbc.co.uk/news/world-europe-48222707.

[4] NCEP/NCAR Reanalysis: Pressure: NOAA Physical Sciences Laboratory. https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.pressure.html.

[5] On Shoulders of Giants | A bid to win 2020 Talisker Whiskey Atlantic Challenge. https://www.onshouldersofgiants.net/. Accessed: 6. 5. 2021.

[6] Talisker Whiskey Atlantic Challenge | The Worlds Toughest Row. https://www.taliskerwhiskyatlanticchallenge.com/. Accessed: 6. 5. 2021.

[7] Edmond Halley. An historical account of the trade winds, and monsoons, observable in the seas between and near the Tropicks, with an attempt to assign the physical cause of the said winds. *Philosophical Transactions of the Royal Society of London*, 16(183):153–168, 1753.

[8] Andy Philpott and Geoff Leyland. Rowing to Barbados. In *Dagstuhl Seminar Proceedings*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2005.