1 **Machine Learning Analysis of Lifeguard Flag Decisions and Recorded Rescues**
2
3 Chris Houser, Jacob Lehner, Nathan Cherry, Phil Wernette
4
5 Faculty of Science,
6 University of Windsor,
7 401 Sunset Avenue,
8 Windsor, Ontario, Canada
9 N9C 2J9
10
11 Corresponding Author: chouser@uwindsor.ca

12 **Abstract**

13

14 Rips currents and other surf hazards are an emerging public health issue globally. Lifeguards,

15 warning flags and signs are important and to varying degrees they are effective strategies to

16 minimize risk to beach users. In the United States and other jurisdictions around the world,

17 lifeguards use coloured flags (green, yellow and red) to indicate whether the danger posed by the

18 surf and rip hazard is low, moderate, or high respectively. The choice of flag depends on the

19 lifeguard(s) monitoring the changing surf conditions along the beach and over the course of the

20 day using both regional surf forecasts and careful observation. There is a potential that the chosen

21 flag is not consistent with the beach user perception of the risk, which may increase the potential

22 for rescues or drownings. In this study, machine learning is used to determine the potential for

23 error in the flags used at Pensacola Beach, and the impact of that error on the number of rescues.

24 Results of a decision tree analysis indicate that the colour flag chosen by the lifeguards was

25 different from what the model predicted for 35% of days between 2004 and 2008 (n=396/1125).

26 Days when there is a difference between the predicted and posted flag colour represent only 17%

27 of all rescue days but those days are associated with ~60% of all rescues between 2004 and 2008.

28 Further analysis reveals that the largest number of rescue days and total number of rescues is

29 associated with days where the flag deployed over-estimated the surf and hazard risk, such as a

30 red or yellow flag flying when the model predicted a green flag would be more appropriate based

31 on the wind and wave forcing alone. While it is possible that the lifeguards were overly cautious

32 it is argued that they most likely identified a rip forced by a transverse-bar and rip morphology

33 common at the study site. Regardless, the results suggest that beach users may be discounting

34 lifeguard warnings if the flag colour is not consistent with how they perceive the surf hazard or

35 the regional forecast. Results suggest that machine learning techniques have the potential to

36 support lifeguards and thereby reduce the number of rescues and drownings.

37

38 **Keywords:** *rip current, surf zone, beach safety, beach hazard*

Deleted: does not accurately reflect
Deleted: the potential risk
Deleted: A
Deleted: suggests
Deleted: that the wrong flag was flown on ~35% of days
Deleted: , and that
Deleted: those differences account for only
Deleted: and
Deleted: the total number of
Deleted: would suggest
Deleted: Regardless whether this is a result of the
Deleted: being
Deleted: or the
Deleted: and surf hazard is associated with weak rips
Deleted: ,
Deleted: are
Deleted: the
Deleted: it isn't

2

Introduction

Rip currents are the main hazard to recreational swimmers and bathers, and, in recent years, have been recognized as a serious global public health issue (Brighton et al., 2013; Woodward et al., 2013; Kumar and Prasad et al., 2014; Arozarena et al., 2015; Brewster et al., 2019; Vlodarchyk et al., 2019). Rips are strong, seaward-directed currents that can develop on beaches characterized by wave breaking within the surf zone (Castelle et al., 2016), and are capable of transporting swimmers a significant distance away from the shoreline into deeper waters. Weak swimmers or those who try and fight the current can become stressed and experience panic (Brander et al., 2011; Drozdzewski et al., 2015) leading to increased adrenaline, an elevated heart rate and blood pressure, and rapid and shallow breathing. On recreational beaches in Australia and the United States, rips have been identified as the main cause of drownings and are believed to be responsible for nearly 80% of all rescues (Brighton et al., 2013; Brewster et al., 2019). It is estimated that the annual number of rip current drownings exceeds the number of fatalities caused by hurricanes, forest fires, and floods in Australia, the United States (Brander et al., 2013; NWS, 2017), while rip-related drownings on a relatively small number of beaches in Costa Rica account for a disproportionately large number of violent deaths in the country (Arozarena et al., 2015). However, recent evidence suggests that public knowledge of this hazard is limited (Brander et al., 2011; Williamson et al., 2011; Brannstrom et al., 2014; 2015; Gallop et al., 2016; Fallon et al., 2018; Menard et al., 2018; Silva-Cavalcanti et al., 2018; Trimble and Houser, 2018), and that few people are interested in rip currents compared to other hazards (Houser et al., 2019).

Many beaches have warning signs at primary access points to warn beach users of the rip hazard, but recent studies suggest that signs may not be effective (e.g. Matthews et al., 2014; Brannstrom et al. 2015). Many beaches also use a combination of beach flags to either designate the location of supervised and safe swimming areas (e.g. Australia and the United Kingdom), or areas and times to avoid entering the water (e.g. Costa Rica and the US). Unfortunately, not every country uses the same flagging convention and there are regional variations that can lead to confusion amongst beach users. The United States and Canada use green, yellow, and red coloured flags to indicate whether the danger posed by the surf and rip hazard is low, moderate, or high, respectively (ILSF, 2004). A beach manager or lifeguard decides on the surf hazard and the flag colour to fly based on a combination of daily updates on rip conditions provided by local lifeguards

Deleted: SLSA, 2017; USLA, 2017

Deleted: , and Costa Rica

Deleted: ;

Deleted: , but

Deleted: in press

as well as a rip forecast from the US National Weather Service (NWS). Most rip forecasts are based on a simple correlation between the number of rip-related rescues and meteorological and oceanographic conditions on that day (Lushine, 1991a, b; Lascody, 1998; Engle, 2002; Dusek and Seim, 2013; Kumar and Prasad, 2014; Scott et al., 2014; Moulton et al., 2017). These forecasts do not account for the surf zone morphology, which may be conducive to the development of rips on days when wave breaking is relatively weak. Even under 'green flag' days, the presence of shore-attached nearshore bars (called a transverse bar and rip morphology; Wright and Short, 1984) can force a current of ~0.5 m s^{-1} that can pose a threat to weak swimmers (Houser et al, 2013).

Rip currents can still be present even if a regional forecast predicts that the hazard potential is low based on wind and wave conditions. Beach users can be at risk if the flag colour is based solely on the regional forecast. To be effective, the flag system requires lifeguards to continuously assess surf conditions and monitor swimmers and bathers, and ultimately intervene if someone does not heed the warning implied by a yellow or red flag indicating moderate and high ('do not enter the water') hazard levels respectively. Recent evidence suggests that many beach users do not adhere to warnings if their own experience (whether accurate or not) or behavior of others on the beach, contradicts the hazard, as indicated by the warning flag (Houser et al., 2017; Menard et al., 2018). Beachgoers may lose trust in authority (i.e. the lifeguards) if a forecast is perceived, wrongly or rightly, to be inaccurate (Espluga et al., 2009). If the forecast is for dangerous surf conditions and a yellow or red flag is placed on the beach when conditions appear to the beach user to be relatively calm, the beach user may discount or ignore the forecast now and, in the future, if they enter the water and do not experience any difficulties. Trust and confidence in the authority figures can be eroded if they believe that the lifeguards are being overly cautious. It can be difficult to change (or 'reset') public perception about the accuracy of the flag system as soon as a discrepancy is perceived, and subsequent visits and experiences may confirm the biases of the beach user (Houser et al., 2018). It is a situation analogous to the boy who cries "*wolf*" (Wachinger et al., 2013).

This study examines the consistency of flag warnings at Pensacola Beach, Florida between 2004 and 2008 when daily data is available for flag colour, wind and wave forcing, as well as the daily number of rescues performed by lifeguards. A decision tree, a form of machine learning, is used to predict the posted flag colour using lifeguard observations in combination with wind and wave forcing. The modelled flag colour, based solely on wave and wind forcing, can be compared

4

Deleted: The presence of a rip when the forecast predicts that the hazard potential is low, can put beach users at risk when a lifeguard is not present and able to intervene/rescue.

Deleted: the warning flag

Deleted: has

Deleted: een

Deleted: and

Deleted: overly-cautious

Formatted: Font: Italic

Formatted: Font color: Auto

Deleted: when daily data is available for flag colour, wind and wave forcing, as well the number of rescues performed by lifeguards

135 to the flag colour posted by the lifeguards on a particular day to identify days when there is a

136 difference and how that influences the number of rescues performed on that day. It is hypothesized

137 that there will be a greater number of rescues performed on days when there is a difference between

138 the predicted and posted flag colour. Specifically, it is hypothesized that a greater number of

139 rescues will occur on days when the model underestimated the hazard level compared to the

140 lifeguard who made their decision based on local observations including the presence of semi-

141 permanent rip channels. In this scenario, the public may believe that the lifeguard is being overly

142 cautious leading to people entering the water.

143

144 **Study Site**

145

146   The analysis was completed at Pensacola Beach, Florida (Figure 1), where there is

147 available records of daily flag colours, wind and wave forcing, and lifeguard-performed rescues

148 between 2004 and 2008. The beaches of the Florida Panhandle have been described ''*as the worst*

149 *in the nation for beach drowning*'' (The Tuscaloosa News, 2002), based on the presence of semi-

150 permanent rips along the length of the island (Houser et al., 2011; Barrett and Houser, 2012). These

151 rips can be active and pose a threat to swimmers when conditions may appear to be safe for

152 swimming (Houser et al., 2013). During the period of the study (2004-2008), the Santa Rosa Island

153 Authority maintained a flagging system to alert beach users about the heavy surf and rip hazard

154 based on the NWS rip forecast. The highest flag colour for that day was recorded by the Island

155 Authority, along with the number of prevents, assists, and rescues. The Island Authority reserve

156 the rescue definition for those persons in extreme difficulty who, in the opinion of the lifeguard,

157 would have drowned without assistance.

158   Rescues, assists, and prevents are recorded regardless of whether they are conducted in a

159 'guarded' area, a designated swimming area where there are typically many beach users (Casino

160 Beach, Fort Pickens Gate Beach, and Park East), or along the ~13 kms of unguarded beach where

161 lifeguards conduct regular patrols and respond to emergency calls. As shown by Barrett and

162 Houser (2013), there are rip current hotspots with semi-permanent alongshore variation in the

163 nearshore morphology due to a ridge and swale bathymetry on the inner shelf. The innermost bar

164 varies alongshore at a scale of ~1000 m, consistent with the ridge and swale bathymetry (Houser

165 et al., 2008), and tends to exhibit a transverse bar and rip morphology immediately landward of

Deleted: posted flag colour

Deleted: discrepancy

Deleted: between the posted and predicted flag colours, which is, in turn, compared to the number o

Deleted: discrepancy

Deleted:

Deleted: for

Deleted:

Deleted: in an

Deleted: inner-shelf

5

176   the deeper swales (Barrett and Houser, 2012; see Figure 1). Historically, most drownings and

177   rescues on this popular beach have occurred at these rip hotspots because they correspond to the

178   main access points along the island (Houser et al., 2015; Trimble and Houser, 2018).

179   Santa Rosa Island experienced widespread erosion and washover during Hurricane Ivan in

180   September 2004. The storm reinforced the alongshore variation in the nearshore bar morphology

181   and forced the bars farther offshore. As described in Houser et al. (2015), the nearshore bars

182   migrated landward and recovered to the beachface for 3 years following the storm. During this

183   period, the inner-bar morphology transitioned from a rhythmic bar and beach morphology to a

184   transverse bar and rip morphology before ultimately attaching to the beachface in May 2008

185   (Houser and Barrett, 2010). This changing bar morphology is a primary control on the presence of

186   rip channels, with the greatest density of rips present in 2005 as the inner-most bar first started to

187   develop a transverse bar and rip morphology (Houser et al., 2011).

188

189   **Methodology**

190

191   Offshore wave conditions and wind forcing function are based on long-term meteorological

192   and oceanographic records from an offshore wave buoy located ~100 km southeast of the study

193   area (buoy 42039; Figure 1). Between 2004 and 2008, this was the closest buoy to Pensacola Beach

194   and had been previously used to estimate the incident wave field (Wang and Horwitz, 2007;

195   Claudino-Sales et al., 2008; 2010; Houser et al., 2011) and was the basis for the rip hazard at

196   Pensacola Beach until a new buoy was placed closer to the beach in 2009. The available wave

197   data from buoy 42039 included offshore significant wave height, significant wave period, and

198   direction, and the wind data included speed and direction. Local water level data was acquired

199   from a station at the Port of Pensacola just north of the study site. A decision tree analysis was

200   used to determine what combination of wave and wind forcing was associated with the flag posted

201   by the Santa Rosa Island Authority on that day. After training on the available dataset, the model

202   produces a decision tree that can be used for future decisions about what flag colour should be

203   posted, although further training would be required to validate the model and operationalize. The

204   modelled (*i.e.* predicted) flag colour is then compared to the posted flag colour for all days to

205   determine if there is a relationship between the flag colour and the number of rescues. The

Deleted: September,

Deleted: s

Deleted: near the study region

Deleted: 42039 and 42040

210   comparison is also used to determine if there is a specific combination of wind and wave forcing

211   on the days when the modelled flag colour and the posted flag colour do not align.

212   A decision tree model was developed using the Chi-square Automatic Interaction Detector

213   (CHAID) technique developed by Kass (1980). The goal of CHAID analysis is to build a model

214   that helps explain how independent variables (wind speed, wave height, wave period, wave

215   direction, wind direction and water level) can be merged to explain the results in a given dependent

216   variable. To develop a decision tree, the first step is declaring the root node, this corresponds to

217   the target variable that will be predicted throughout the model. Then, the independent variable that

218   provides the most information about the target values is identified. The root node is then split on

219   this independent variable into statistically significant different subgroups using the F-test. These

220   subgroups are then split using the predictor variables that provide the most information about them.

221   CHAID analysis continues this process until terminal nodes are reached and no splits are

222   statistically significant. Previous use of CHAID analysis in hazard studies include landslide

223   prediction (e.g. Althuwaynee et al., 2014), farmer perception of flooding hazard (Bielders et al.,

224   2003; Tehrany et al., 2015), and property owner perception and decision making along an eroding

225   coast (Smith et al., 2017).

226

227   **Results**

228

229   The decision tree model was trained on the 1125 days with complete data between 2004

230   and 2008. Over this same period there were 145 days with rescues. The annual number of rescues

231   and rescue days (ie. days with one or more rescues) varied by year, with a peak in both the total

232   number of rescues and the number of rescue days in 2005. The number of rescues was at a

233   minimum in 2007, while the number of rescue days was at a minimum in 2006 (Figure 3). The

234   number of rescues decreased linearly between 2005 and 2007 as the nearshore bar morphology

235   continued to recover following Hurricane Ivan and welded to the beachface consistent with

236   previous observations at the site (Houser et al., 2011). It is important to note that the CHAID

237   Analysis does not incorporate nearshore morphology as an independent variable because changes

238   in nearshore morphology were not tracked daily over the study period. In this respect, differences

239   between the posted and predicted flag colour may reflect lifeguard observations of nearshore

Deleted: The

Deleted: .

Deleted: during which there was

243 morphology conducive to the development of rip currents despite winds and waves typical of green
244 flag conditions.

245   The decision tree analysis suggests that the posted flag colour was not predicted by the
246 model on 35% of days between 2004 and 2008 (n=396). There was a total of 342 rescues over 66
247 days when the model predicted a different flag than was posted representing over 60% of all
248 rescues (Table 1). By comparison, 40% of all rescues (n=224) occurred over 79 days when the
249 predicted and posted flags were the same. Chi-square analysis suggests that the number of rescue
250 days is significantly greater at the 95% confidence level when the predicted and posted flags are
251 different (χ^2=7.77, ρ~0.005). This supports the hypothesis that there are a greater number of
252 rescues performed on days when there is a discrepancy between the predicted and posted flag
253 colour.

254

255 **Table 1.** Results of Chi-square analysis of posted and predicted flag colour versus rescue and no
256 rescue days at Pensacola Beach, Florida between 2004 and 2008.
257

	Rescue Days	No Rescue Days	
Posted=Predicted	79	650	χ^2=7.77, ρ~0.005
Posted≠Predicted	66	330	

258

259   Chi-square analysis was also used to determine if the number of rescue days depends on
260 whether the model predicts a flag of greater or lesser hazard compared to the posted flag (Table
261 2). Results suggest that the number of rescue days is greater when the model predicts hazardous
262 surf (i.e. red or yellow flag), but the posted flag was either yellow or green (χ^2=18.11, ρ~0.0001).
263 The number of rescue days was over-represented when the posted flag colour was red or yellow,
264 but the model predicted that the flag should have been yellow or green, respectively, suggesting
265 that posting what a beach user may perceive as an overly cautious flag can present a danger. These
266 47 days were associated with 268 of the total 566 rescues between 2004 and 2008, or ~7.2 rescues
267 per day when the island authority posted a more cautious flag then was predicted by the model.
268 In comparison, the number of rescues (n=298) was under-represented on days when the posted
269 flag suggested conditions were not as hazardous (n=74) as the model or were identical to the model
270 (n=224).

271

Deleted: primary hypothesis that
Deleted: will be
Deleted: a
Deleted: overly-cautious
Deleted: was
Deleted: overly-cautious in their flag choice

278 **Table 2.** Results of Chi-square analysis of posted and predicted flag colour versus rescue and no
279 rescue days at Pensacola Beach, Florida between 2004 and 2008.
280

	Rescue Days	No Rescue Days	
Posted>Predicted	47	171	χ^2=18.11, ρ~0.0001
Posted<Predicted	19	159	
Posted=Predicted	79	650	

281

282 The greatest number of rescues were performed on days when the posted flag was yellow
283 (moderate hazard, moderate surf and/or currents), but the model predicted a green flag (low hazard,
284 relatively calm surf and/or currents) based on the wind and wave forcing. Specifically, a total of
285 231 rescues were performed on 37 of the 168 days when the posted flag was yellow, and the model
286 predicted that the flag colour should be green. In comparison, there were only 12 rescues on 3 of
287 20 days when the posted flag was red (high hazard, strong surf and/or currents) and the model
288 predicted flag colour was green. Finally, there were 25 rescues preformed on 7 of 30 days when
289 a red flag was posted, and the model predicted a yellow flag was appropriate. The number of
290 rescues and rescue days when the posted flag was more cautious than predicted by the model were
291 at a maximum in 2005 and linearly decreased to a minimum in 2007 as the bar morphology
292 recovered from Hurricane Ivan.

293 While there were fewer than expected rescue days when the posted flag was green or
294 yellow and the model predicted a yellow or red flag, rescues were still performed on those days.
295 There was a total of 66 rescues on 13 of 80 days when the posted flag was yellow, but the model
296 predicted a red flag should be posted (Table 3). Only 7 rescues were performed on 5 of the 83 days
297 when the posted flag was green and the model predicted a yellow flag, with even fewer rescues
298 performed on days when the posted flag was green, but should have been red. The number of
299 rescues and rescue days when the posted flag was lower than the predicted flag decreased from
300 2004 to 2007, with a statistically significant outlier in 2008. The large number of rescues in 2008
301 is the result of 2 days with 13 rescues each (April 19 and September 14), when a yellow flag was
302 being flown, but the model predicted a red flag was more appropriate. This suggests that the
303 difference between posted and predicted flag colours can vary inter-annually with changes in the
304 nearshore morphology and/or changes in the individual who makes the flag decision.
305

9

Table 3. Number of days and rescues (in brackets) based on the combination of posted and predicted flag colours.

		Predicted Flag		
		G	**Y**	**R**
Posted Flag	**G**	475 (48)	83 (7)	15 (1)
	Y	168 (231)	154 (125)	80 (66)
	R	20 (12)	30 (25)	100 (51)

Discussion

Results of the present study suggest that over 60% of all rescues at Pensacola Beach, Florida between 2004 and 2008 occurred on days when the posted hazard flag was different from the flag colour predicted by a decision tree model. The posted flag colour was not predicted by the model on 35% of days between 2004 and 2008 (n=396), with one or more rescues occurring on 66 of those days (~17%). While rescues did not occur on a vast majority of the days when the posted and predicted flag colours were different, days when the predicted and posted flag colours were different accounted for a majority of the rescues. This is not to suggest that Santa Island Authority made a mistake in their flag choice. Rather, the results suggest that the difference between the posted and predicted flag colour could be associated with the lifeguards noting that the nearshore had a transverse bar and rip morphology, which is common at this location. The morphology of the nearshore and other variables that could influence whether a beach user will enter the water or not (e.g. weather, number of beach users or presence of seaweed) are not captured by the current model, which is based on wind and wave forcing alone. The model developed in this study is similar to rip forecasts produced by the US National Weather Service (NWS), and does not include local variables known to the beach manager based on experience and years of careful observation. Discrepancies between the predicted and posted flag colours provide a basis for future model development and expansion. Incorporating more data into the model will it to evolve and better capture the variables that influence the colour of flag chosen by the lifeguards, while ensuring that the model remains computationally efficient. Introducing additional variables, such as nearshore morphology, to the model has the potential to better capture a lifeguard or beach manager's understanding of what constitutes dangerous surf conditions at their beach. At the same time, it is also important to examine the accuracy of beach managers and lifeguards in assessing the nearshore morphology and potential for rip development.

366   The model predictions and most forecasts are based solely on wind and wave forcing
367   (Lushine, 1991a, b; Lascody, 1998; Engle, 2002; Dusek and Seim 2013; Arun Kumar and Prasad,
368   2014; Scott et al., 2014; Moulton et al., 2017). Noticeably absent from the current model is surf
369   zone morphology, which ultimately determines whether a rip can develop under those conditions
370   or not. The beach manager and lifeguard can observe the nearshore morphology and assess the
371   potential for rip development, which would lead to them putting out a yellow or red flag when the
372   model would predict a green or yellow flag as being appropriate. While beach managers and
373   lifeguards are being prudent, their assessment may not conform to those of the beach user who
374   decides on whether the water is safe or not based on wave breaking conditions (Caldwell et al.,
375   2013; Brannstrom et al., 2013; 2015). Most beach users assume that larger breaking waves are
376   more dangerous, and many will not enter the water if they (and the model) believe that it is a 'red'
377   flag condition. This may partially explain why there were fewer than expected rescues on days
378   when the posted flag colour was green or yellow flag and the model predicted a yellow or red flag,
379   respectively. Independent of the flag or warning signs, beach users appear to be making personal
380   decisions about the surf and rip hazard (Brannstrom et al., 2015) based on experience at the site or
381   elsewhere (see Houser et al., 2018). Whether this causes beach users to lose confidence in the
382   lifeguards and other authorities managing the beach is an important question for future research.
383   A large number of rescues occurred when the posted flag was yellow, but the model
384   predicted the wind and wave forcing warranted a green flag. Rightly or wrongly, the beach user
385   will observe that wave breaking is limited and assume that conditions must be safe. As shown by
386   Caldwell et al. (2013) and Brannstrom et al. (2013) most beach users along the Gulf Coast of the
387   United States assume that the calm flat water of a rip is safer than adjacent areas where the waves
388   are breaking. The lifeguard, however, may observe a bar morphology that is conducive to the
389   development of rips and post a yellow flag to warn about the potential for rips, despite the weak
390   wind and wave forcing. As observed by Houser and Barrett (2012), rips with speeds of ~0.5 m/s
391   can develop on 'green flag' days because of the transverse bar and rip morphology that is present
392   in the inner-nearshore. This would suggest that posting a green flag should never be permitted
393   when wind and swell waves are breaking over the bar, even if the regional forecast suggests a low-
394   level hazard that day. As shown by Scott et al. (2014), rescues are still possible with seemingly
395   'fine weather' conditions when a green flag would be predicted by the model or in regional

Deleted: is

Deleted: conditions

Deleted: was overly-conservative (*e.g.*

Deleted: was

Deleted: posted when the

Deleted:)

Deleted: erodes

Deleted: '

404 forecasts. Even in the presence of small swell wave, breaking can be induced as water levels fall

405 with the tide (Castelle et al. 2016).

406   It is difficult for beach users to spot a rip or assess the potential for rip development, and

407 they may assume that the lifeguard is being overly cautious if they perceive fine-weather

408 conditions and the lifeguard posts a yellow or red flag. Going to the beach is a reward-based

409 activity, and many people commit significant personal and financial investment to be at the beach

410 (Houser et al., 2018). If they believe that the lifeguard is 'wrong' they will ignore the warning and

411 remain committed to entering the water. The longer and more times that their perceptions are

412 inconsistent with the experience and knowledge of the lifeguard, the more trust in authority is lost

413 - a beach that is perceived to be safe based on experience will always be safe despite warnings to

414 the contrary (Menard et al., 2018). This is an example of confirmation bias, in which an opinion

415 quickly becomes entrenched and subsequent evidence is used to either bolster the belief or is

416 rapidly discarded. How this can be addressed to reduce the number of rescues is an important focus

417 for future research on rips and other hazards in general.

418   The results of this study also highlight the limitations of regional rip forecasts that are used

419 in the United States and elsewhere around the world. A forecast based solely on the wind and wave

420 forcing does not account for the nearshore morphology, which determines the potential for rip

421 development. This raises one of the most important considerations for future modeling efforts

422 based on machine learning techniques - the model will only be accurate if the bar morphology and

423 conceptual knowledge of the lifeguard is included as input variables. Getting the beach user to

424 observe and heed that forecast and warning, however, will remain a challenge.

425

426 **Conclusions**

427   Lifeguards and beach managers decide on warnings and flag colours based on careful

428 monitoring of the changing surf conditions along the beach and over the course of the day using

429 both regional surf forecasts and direct observation. A decision tree analysis predicts a flag colour

430 different to the one flown on ~35% of days between 2004 and 2008 (n=396/1125), and that those

431 differences account for only 17% of all rescue days and ~60% of the total number of rescues. The

432 posting of a yellow flag when the model would predict a green flag based solely on the wind and

433 wave forcing was found to be responsible for the largest number of rescues over the study period.

434 Variables such as the nearshore morphology and the potential for rip development is not included

Deleted: .

Deleted: eroded

Deleted: the

Deleted: The

439 in traditional forecasts or the model developed in this paper, and most beach users use a simple
440 assessment of wave breaking to determine if the water is safe. Even though a lifeguard will post
441 the appropriate flag based on direct observation of the bar morphology and experience, the beach
442 user, like simple models based solely on meteorological data, may not believe that warning and
443 still enter the water. This suggests that reducing the number of rip and surf rescues will require
444 that we are able to address confirmation bias on the part of the beach user, which can cause them
445 to lose their confidence in the lifeguards.
446
447 **Acknowledgements**
448
449 This study was partly funded through a NSERC Discovery Grant to CH.

Deleted: erode their confidence

13

References

451 **References**

453 Althuwaynee, O.F., Pradhan, B., Park, H.J. and Lee, J.H.. A novel ensemble decision tree-based

454   CHi-squared Automatic Interaction Detection (CHAID) and multivariate logistic

455   regression models in landslide susceptibility mapping. *Landslides*, *11*(6), 1063-1078,

456   2014.

457 Arozarena, I., Houser, C., Echeverria, A.G., and Brannstrom, C.: The rip current hazard in Costa

458   Rica, Natural Hazards, 77(2), 753-768, 2015.

459 Barrett, G. and Houser, C. Identifying hotspots of rip current activity using wavelet analysis at

460   Pensacola Beach, Florida. *Physical Geography*, *33*(1), 32-49, 2012.

461 Bielders, C.L., Ramelot, C. and Persoons, E. Farmer perception of runoff and erosion and extent

462   of flooding in the silt-loam belt of the Belgian Walloon Region. *Environmental Science*

463   *& Policy*, *6*(1), 85-93, 2003.

464 Brander, R.W., Bradstreet, A., Sherker, S., and MacMahan, J.: Responses of swimmers caught in

465   rip currents: Perspectives on mitigating the global rip current hazard, International

466   Journal of Aquatic Research and Education, 5(4), 11, 2011.

467 Brander, R.W., Dominey-Howes, D., Champion, C., Del Vecchio, O., and Brighton, B.: Brief

468   Communication: A new perspective on the Australian rip current hazard, Natural Hazards

469   and Earth System Sciences, 13(6), 1687, 2013.

470 Brannstrom, C., Trimble, S., Santos, A., Brown, H.L., and Houser, C.: Perception of the rip

471   current hazard on Galveston Island and North Padre Island, Texas, USA, Natural

472   Hazards. 72(2), 1123-1138, 2014.

14

473 Brannstrom, C., Brown, H.L., Houser, C., Trimble, S., and Santos, A.: "You can't see them from
474   sitting here": Evaluating beach user understanding of a rip current warning sign, Applied
475   Geography, 56, 61-70, 2015.

476 Brewster, B.C., Gould, R.E. and Brander, R.W. Estimations of rip current rescues and drowning
477   in the United States. *Natural Hazards and Earth System Sciences*, *19*(2), 389-397, 2019.

478 Brighton, B., Sherker, S., Brander, R., Thompson, M. and Bradstreet, A. Rip current related
479   drowning deaths and rescues in Australia 2004–2011. *Natural hazards and earth system*
480   *sciences*, *13*(4), pp.1069-1075, 2013.

481 Caldwell, N., Houser, C., and Meyer-Arendt, K.: Ability of beach users to identify rip currents at
482   Pensacola Beach, Florida, Natural Hazards, 68(2), 1041-1056, 2013.

483 Castelle, B., Scott, T., Brander, R., and McCarroll, R.J.: Rip current types, circulation and
484   hazard, Earth-Science Reviews, 163, 1-21, 2016.

485 Claudino-Sales, V., Wang, P. and Horwitz, M.H. Factors controlling the survival of coastal dunes
486   during multiple hurricane impacts in 2004 and 2005: Santa Rosa barrier island,
487   Florida. *Geomorphology*, *95*(3-4), 295-315, 2008.

488 Claudino-Sales, V., Wang, P. and Horwitz, M.H. Effect of Hurricane Ivan on coastal dunes of Santa
489   Rosa Barrier Island, Florida: characterized on the basis of pre-and poststorm LIDAR
490   surveys. *Journal of Coastal Research*, 470-484, 2010.

491 Drozdzewski, D., Shaw, W., Dominey-Howes, D., Brander, R., Walton, T., Gero, A., and
492   Edwick, B.: Surveying rip current survivors: preliminary insights into the experiences of
493   being caught in rip currents, Natural Hazards and Earth System Sciences, 12(4), 1201,
494   2012.

15

Dusek, G., and Seim, H.: A probabilistic rip current forecast model, Journal of Coastal Research
      29(4), 909-925, 2013.

Engle, J., Macmahan, J., Thieke, R. J., Hanes, D. M., and Dean, R.G.: Formulation of a Rip
      Current Predictive Index Using Rescue Data. In *Proc. National Conf. on Beach*
      Preservation Technology, FSBPA, 2012.

Espluga, J., Farré, J., Gonzalo, J., Horlick-Jones, T., Prades, A., Oltra, C., and Navajas, J.: Do
      the people exposed to a technological risk always want more information about it? Some
      observations on cases of rejection. Safety, Reliability and Risk Analysis. Londres: CRC
      Press. Taylor & Francis, 1301-1308, 2009.

Fallon, K.M., Lai, Q. and Leatherman, S.P. Rip current literacy of beachgoers at Miami Beach,
      Florida. *Natural hazards*, *90*(2), 601-621, 2018.

Gallop, S.L., Woodward, E., Brander, R.W. and Pitman, S.J. Perceptions of rip current myths
      from the central south coast of England. *Ocean & coastal management*, *119*, 14-20, 2016.

Houser, C., Barrett, G., and Labude, D.: Alongshore variation in the rip current hazard at
      Pensacola Beach, Florida, Natural Hazards, 57(2), 501-523, 2011.

Houser, C., Arnott, R., Ulzhöfer, S., and Barrett, G.: Nearshore Circulation over Transverse Bar
      and Rip Morphology with Oblique Wave Forcing, Earth Surface Processes and
      Landforms, 38, 1269-1279, 2013.

Houser, C., Hapke, C. and Hamilton, S. Controls on coastal dune morphology, shoreline erosion
      and barrier island response to extreme storms. *Geomorphology*, *100*(3-4), 223-240, 2008.

Houser, C., Wernette, P., Rentschler, T., Jones, H., Hammond, B. and Trimble, S.: Post-storm
      beach and dune recovery: implications for barrier island resilience. Geomorphology, 234,
      54-63, 2015.

Deleted: resilience, Geomorphology

16

519 Houser, C., and Barrett, G.: Divergent behavior of the swash zone in response to different

520   foreshore slopes and nearshore states, Marine Geology, 271, 106-118, 2010.

521 Houser, C., Murphy, T., and Labude, D.: Alongshore correspondence of beach users and rip

522   channels at Pensacola Beach, Florida, Natural Hazards, 78(3), 2175-2183, 2015.

523 Houser, C., Trimble, S., Brander, R., Brewster, B.C., Dusek, G., Jones, D., and Kuhn, J.: Public

524   perceptions of a rip current hazard education program: "Break the Grip of the

525   Rip!", Natural hazards and earth system sciences, 17(7), 1003, 2017.

526 Houser, C., Vlodarchyk, B. and Wernette, P. public interest in rip currents relative to other

527   natural hazards: evidence from Google search data. *Natural Hazards*, 1-11, 2019.

528 Kass, G.V.: An exploratory technique for investigating large quantities of categorical data,

529   Applied Statistics, 29(2), 119-127, 1980.

530 Kumar, S.A., and Prasad, K.V.S.R.: Rip current-related fatalities in India: a new predictive risk

531   scale for forecasting rip currents, Natural hazards, 70(1), 313-335, 2014.

532 Lascody, R.L.: East central Florida rip current program, National Weather Digest, 22(2), 25-30,

533   1998.

534 Lushine, J.B.: A study of rip current drownings and related weather factors, In Natl. Wea. Dig,

535   1991.

536 Lushine, J.B.: Rip Currents: Human Impact and Forecastibility, In Coastal Zone'91, 3558-3569.

537   ASCE, 1991.

538 Matthews, B., Andronaco, R., and Adams, A.: Warning signs at beaches: Do they work? Safety

539   science, 62, 312-318, 2014

Ménard, A.D., Houser, C., Brander, R.W., Trimble, S. and Scaman, A. The psychology of beach
users: importance of confirmation bias, action, and intention to improving rip current
safety. *Natural Hazards*, *94*(2), 953-973, 2018.

Moulton, M., Dusek, G., Elgar, S., and Raubenheimer, B.: Comparison of Rip Current Hazard
Likelihood Forecasts with Observed Rip Current Speeds, Weather and Forecasting,
32(4), 1659-1666, 2017.

Scott, T., Masselink, G., Austin, M.J., and Russell, P.: Controls on macrotidal rip current
circulation and hazard, Geomorphology, 214, 198-215, 2014.

Silva-Cavalcanti, J.S., Costa, M.F. and Pereira, P.S. Rip currents signaling and users behaviour
at an overcrowded urban beach. *Ocean & coastal management*, *155*, 90-97, 2018.

Surf Life Saving Australia (SLSA): National Coastal Safety Report 2017. Surf Life Saving
Australia, Sydney, 2017.

Tehrany, M.S., Pradhan, B. and Jebur, M.N. Flood susceptibility analysis and its verification
using a novel ensemble support vector machine and frequency ratio method. *Stochastic
environmental research and risk assessment*, *29*(4), 1149-1165, 2015.

Trimble, S., and Houser, C.: Seawalls and Signage: How Beach Access Management Affects Rip
Current Safety. In Beach Management Tools-Concepts, Methodologies and Case Studies,
497-524, Springer, Cham, 2017.

Vlodarchyk, B., Olivito, A. and Houser, C. Spatial and Temporal Variation of Surf Drownings in
the Great Lakes: 2010–17. *Journal of Coastal Research*, *35*(4), 794-804, 2019.

Wachinger, G., Renn, O., Begg, C., Kuhlicke, C.: The risk perception paradox—implications for
governance and communication of natural hazards, Risk analysis, 33(6), 1049-1065,
2013.

565   Wang, P. and Horwitz, M.H. Erosional and depositional characteristics of regional overwash

566   deposits caused by multiple hurricanes. *Sedimentology*, *54*(3), 545-564, 2007.

567   Williamson, A., Hatfield, J., Sherker, S., Brander, R. and Hayen, A. A comparison of attitudes

568   and knowledge of beach safety in Australia for beachgoers, rural residents and

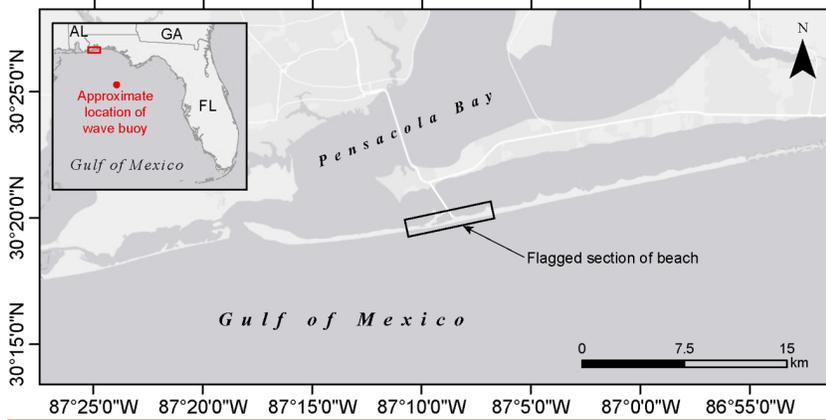569   international tourists. *Australian and New Zealand Journal of Public Health*, *36*(4), 385-

570   391, 2012.

571   Woodward, E., Beaumont, E., Russell, P., Wooler, A. and Macleod, R. Analysis of rip current

572   incidents and victim demographics in the UK. *Journal of Coastal Research*, *65*(sp1),

573   850-856, 2013.

574   Wright, L.D. and Short, A.D. Morphodynamic variability of surf zones and beaches: a

575   synthesis. *Marine geology*, *56*(1-4), 93-118, 1984.

576 **Figures**

577

578

579

580 **Figure 1.** Map of study site showing location of flagged section of beach and approximate
581 location of the wave buoy used in the analysis and for regional rip forecasts.
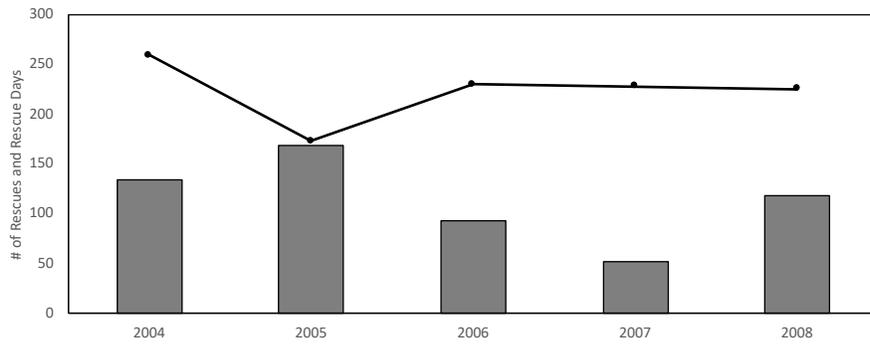

582

592

593

594 **Figure 2.** Satellite image of the flagged section of beach in 2004 (before Hurricane Ivan)
595 showing the presence of transverse-bar and rip morphology of the innermost bar and the variable
596 nature of the outermost bar for the flagged section of beach. The aerial image is not necessarily
597 representative of the nearshore morphology throughout the remainder of the study.
598

599

Figure 3. Interannual variation in number of rescues and rescue days at Pensacola Beach between 2004 and 2008.